INTERNAL ATMOSPHERIC GRAVITY WAVES AT IONOSPHERIC HEIGHTS

C. O. HINES

Reprinted from
CANADIAN JOURNAL OF PHYSICS
38, 1441 (1960)

INTERNAL ATMOSPHERIC GRAVITY WAVES AT IONOSPHERIC HEIGHTS[1]

C. O. HINES

ABSTRACT

Irregularities and irregular motions in the upper atmosphere have been detected and studied by a variety of techniques during recent years, but their proper interpretation has yet to be established. It is shown here that many or most of the observational data may be interpreted on the basis of a single physical mechanism, namely, internal atmospheric gravity waves.

A comprehensive picture is envisaged for the motions normally encountered, in which a spectrum of waves is generated at low levels of the atmosphere and propagated upwards. The propagational effects of amplification, reflection, intermodulation, and dissipation act to change the spectrum continuously with increasing height, and so produce different types of dominant modes at different heights. These changes, coupled with an observational selection in some cases, lead to the various characteristics revealed by the different observing techniques. The generation of abnormal waves locally in the ionosphere appears to be possible, and it seems able to account for unusual motions sometimes observed.

I. INTRODUCTION

Much attention has been devoted in recent years to the detection and measurement of irregular motions in the D, E, and lower F regions of the upper atmosphere, and to the occurrence of irregular density distributions at the same heights. Only tentative interpretations have been put forward until recently, and in the main these have been based on the presumed occurrence of turbulence. It appears, however, that many of the motions and inhomogeneities may have their origin in a much more organized fluctuation of the atmosphere, namely, in propagating atmospheric waves controlled by gravitational and compressional forces. The preliminary support for such a conclusion has been presented elsewhere (Hines 1959*a*); it is here amplified and in some respects superseded.

Meteor trails provide the most direct evidence for motions of the type to be considered, and their observation has yielded the best quantitative data now available for the heights at which they lie. Some characteristics of the motions they reveal are summarized in Section II, and are there compared with the corresponding properties of internal atmospheric gravity waves. These waves are found to provide an adequate interpretation of the observations. Two possible sources of their energy are proposed and discussed: tidal oscillations and low-altitude wind systems.

Some other manifestations of irregularities and irregular motions at very similar heights are discussed in Section III. They consist of partial reflection and scattering of radio waves from inhomogeneities of ionization in the D and E regions, and 'drift' observations of radio diffraction patterns produced by moving irregularities in the E and lower F regions. The possible role of

[1]Manuscript received July 7, 1960.
Contribution from the D.R.B. Theoretical Studies Group, Defence Research Board, Ottawa, Canada.

atmospheric waves in each of these phenomena is indicated but not pursued in detail. A relatively distinct form of moving irregularity, most often detected in the *F* region and usually termed a 'large-scale travelling disturbance', seems likely to find its most appropriate interpretation in terms of the same wave mechanism, and it, too, is discussed in Section III.

The mathematical development is confined to Section IV, and is in the nature of an appendix to the earlier arguments. The analysis is simplified to an extreme, for it largely neglects certain aspects of the actual wave process which must be taken into account in any comprehensive development. This simplified attack is useful, however, in that it permits a straightforward derivation and hence a clearer appreciation of the essential features of what has been until now a relatively neglected type of wave motion. The complications encountered in the actual atmosphere will change some of the deduced relations quantitatively and in certain respects qualitatively, but with one exception they are set aside for subsequent study.

Some concluding remarks are contained in Section V.

II. WINDS REVEALED BY THE DISTORTION OF METEOR TRAINS

II.1. Outline of Observations

Long-enduring meteor trains often become rapidly distorted from their initial straight-line shape, being carried about by 'winds' in the atmospheric regions where they are formed (roughly 80–115 km above the earth). Early visual means of determining the winds have now given way to much more accurate photographic observations (notably, those made under the auspices of the Harvard College Observatory) and radar techniques (applied most thoroughly at the University of Manchester, Stanford University, and the University of Adelaide).

No attempt will be made here to summarize all aspects of these observations nor the details of their earlier interpretation. Instead, some of the salient features pertaining mainly to the smaller scale irregular motions will be recorded, in order to permit a rapid comparison with the wave interpretation. These features are abstracted primarily from the papers of Liller and Whipple (1954) and Greenhow and Neufeld (1959), but other publications and private discussions support a belief that they are fairly representative of the motions concerned.

The winds under discussion are illustrated by Fig. 1, which is drawn after Liller and Whipple (1954). The horizontal and vertical scales of this figure are such that the depicted curve could represent a direct photograph of a distorted meteor train, 200 seconds after the formation of the train along a vertical straight line. In actual fact, the velocities were deduced from a succession of photographs extending over a period of only 17 seconds.

A large-scale shear of the wind system is apparent in Fig. 1, changing from a general motion towards the right at the bottom into a general motion towards the left at the top. This shear is probably attributable to prevailing and tidal winds; it is not of immediate concern here.

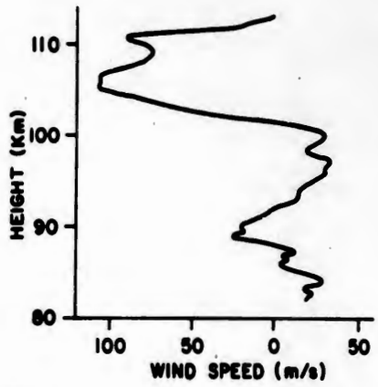

FIG. 1. Sample profile of one component of the wind velocity at meteor heights, after Liller and Whipple (1954). The abscissa measures speed to the right and to the left from the '0' position; speeds to the left are treated as negative in the text.

Superimposed on the general shear of the wind system are very strong smaller scale fluctuations, characterized by reversals in vertical distances of a few kilometers and revealing winds of the order of tens of meters per second. It is these irregular winds which are here to be interpreted in terms of atmospheric waves. They are to be thought of as random movements resulting from the superpositioning of many oscillatory motions, each in turn associated with a specific mode of atmospheric wave propagation.

The salient properties of these irregular winds, which any theoretical interpretation must seek to explain, include the following:

(i) The irregular wind components exhibit strong variations in vertical distances of a few kilometers. The 90-km level may be taken as representative, and there a height interval of about 6 km is sufficient to provide a complete reversal of the irregular velocity (Millman 1959; Bowles 1959). This is about half the interval that might be implied from autocorrelation analyses, wherein the autocorrelation coefficient falls to zero and not to a negative extremum for vertical separations of 6 km (Liller and Whipple 1954; Greenhow and Neufeld 1959), but it seems to be more appropriate in the present context. The discrepancy is associated with a contamination of the autocorrelation values, introduced by the large-scale shear and by the dominating winds of the higher altitudes, as will be seen quantitatively later in this section. The 6-km scale of reversal leads to a value of 12 km for the 'vertical wavelength' ($\equiv$ the vertical spacing between two nodal planes whose phases differ by 2π) in the wave interpretation, and this value will be adopted accordingly as representative of the dominant mode of oscillation.

(ii) The dominant winds at any given position persist with little change for many tens of minutes at a time, and tend to reach zero autocorrelation only after an interval of about 100 minutes (Greenhow and Neufeld 1959). The scaling factor which relates this value to the period of the dominant mode, in the wave interpretation, is not at all clear, if only because contamination by unresolved tidal components may be present, but an empirical analogy with

the results quoted in (i) suggests a value of 200 minutes as a representative period. This value is certainly appropriate within a factor of 2, and such a factor is shown explicitly in the subsequent illustrations. (The theory should, however, include coriolis forces at periods longer than 200 minutes, and this complication has not been taken into account.)

(iii) The horizontal scale size of the irregular winds exceeds the dominant vertical scale size by a factor of 20 or more (Greenhow and Neufeld 1959). The most representative value to employ as a 'horizontal wavelength' ($\equiv$ the horizontal spacing between two nodal planes whose phases differ by 2π) is subject to uncertainties associated with the autocorrelation analysis, as before, but it must lie in the hundreds of kilometers.

(iv) The dominant motions are nearly horizontal. Vertical motions have been sought without substantial success, with the possible exception of the smallest scale components (Manning, Peterson, and Villard 1954; Manning 1959). Most observational data are now routinely reduced on the assumption that the motion is essentially horizontal at meteor heights. In terms appropriate to a wave interpretation, it seems reasonable to assert that any vertical motions spectrally associated with the dominant horizontal motions are less than 10% of the latter.

(v) The speed of the dominant irregular winds tends to increase with height (Liller and Whipple 1954).

(vi) The dominant scale size increases with height, as may be seen in Fig. 1 and in other curves presented by Liller and Whipple. The smallest detected vertical structure size similarly increases with height. Vertical wavelengths of 1 km centered at a height of 87 km, and 6 km centered at a height of 108 km, are typical of the smallest scales depicted in Fig. 1, and may be indicative of general conditions. (The 1-km value might represent a minimum resolution in the optical technique, but Manning (1959) has derived it by quite independent radio means.)

In the remainder of this section, it will become apparent that properties (iii)–(vi) can be deduced directly from the assumption of atmospheric waves as the pertinent mechanism, once properties (i) and (ii) are specified. The basic causes of properties (i) and (ii) may be expected to be found when the excitation process is understood, and first attempts at such an understanding are therefore also indicated.

II.2. The Wave Interpretation

The properties of adiabatic waves propagating through an otherwise stationary and isothermal atmosphere are described in Section IV (which might best be read at this point, by those who prefer to establish the mathematical background before discussing its implications). Both compressional and gravitational forces must be considered in the treatment, and at periods of many minutes the gravitational forces introduce a profound anisotropy into the behavior pattern. The waves of present concern, with periods of 200 minutes or so, may then be described conveniently as atmospheric gravity waves.

Two distinct classes of these waves merit special consideration. They share the property that they are unattenuated in the horizontal direction (in the absence of dissipative forces), but they differ radically in their vertical variations. The one class, which may be designated as 'surface' waves, may vary exponentially in the vertical direction but can support no phase propagation in that direction. Specifically, these waves are unable to provide the vertical phase variations which would account for the vertical structure observed. They will therefore receive no further consideration here.

The second class consists of 'internal' waves, which can support a substantial vertical component of phase propagation. These waves are the low-frequency equivalents of acoustic waves, and they include, if coriolis forces are taken into account, the tidal waves which are already known to exist in the atmosphere. Their properties will now be examined with respect to the irregular meteor-height winds.

Internal atmospheric gravity waves exist at periods greater than a certain characteristic value, given by

$$\tau_g = 2\pi C/g(\gamma-1)^{\frac{1}{2}} \tag{1}$$

in the idealized circumstances treated in Section IV. Here C is the speed of sound ($\simeq 280$ m/s, say, at meteor heights), g is the gravitational acceleration ($\simeq 9.5$ m/s^2 in the same region), and γ is the ratio of specific heats (probably 1.40). Thus $\tau_g \simeq 4.9$ minutes in the present instance. It is evident that the periods of the dominant observed modes greatly exceed τ_g.

In these circumstances, and for sufficiently small structure sizes (including those revealed by the deformation of meteor trains), certain 'asymptotic' relations apply to the parameters that describe these waves. One such is the asymptotic dispersion relation, which may be written in the form

$$\lambda_x/\lambda_z \simeq \tau/\tau_g, \tag{2}$$

where λ_x is the horizontal wavelength, λ_z is the vertical wavelength, and τ is the period of the oscillation. (The mathematical analysis of Section IV employs the standard form of wave-number components, k_x and k_z; in terms of them, $\lambda_x \equiv |2\pi/k_x|$ and $\lambda_z \equiv |2\pi/k_z|$.) Insertion of the values $\lambda_z \simeq 12$ km and $\tau \simeq 200$ minutes, from properties (i) and (ii) above, then yields $\lambda_x \simeq 490$ km which is certainly adequate to meet property (iii).

A second applicable asymptotic relation is that connecting the horizontal and vertical components of atmospheric velocity, U_x and U_z:

$$U_x/U_z \simeq \lambda_x/\lambda_z. \tag{3}$$

Insertion of the preceding values will now reveal that U_x exceeds U_z by a factor of 40, which is in accord with property (iv).

Internal atmospheric gravity waves tend to increase their amplitudes of oscillation with increasing height, in proportion to $\exp(\gamma gz/2C^2)$ where z is a co-ordinate measured vertically upwards. This variation is readily interpreted in terms of energy flux, since it just compensates for the upward decrease of atmospheric density (in proportion to $\exp(-\gamma gz/C^2)$) in maintaining the flux

constant. The amplification is of great significance to the whole subject of these waves, as will be seen in the following subsection. Here, however, it may simply be noted that the amplification appears to account for property (v) listed above, although a quantitative comparison is complicated by superimposed damping effects due to energy dissipation, and may be further complicated by intermodulation. In the absence of such effects, Fig. 2 gives a pictorial summary of the various attributes of the dominant modes as discussed so far.

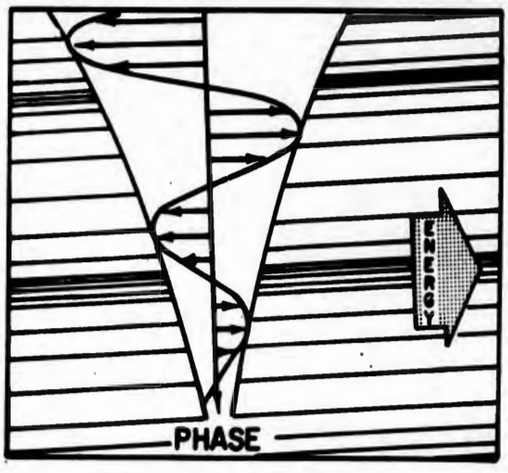

FIG. 2. Pictorial representation of internal atmospheric gravity waves. Instantaneous velocity vectors are shown, together with their instantaneous and over-all envelopes. Density variations are depicted by a background of parallel lines lying in surfaces of constant phase. Phase progression is essentially downwards in this case, and energy propagation obliquely upwards; gravity is directed vertically downwards.

Two basic causes of energy dissipation necessarily affect wave propagation: molecular viscosity and thermal conductivity. These generally become significant under very similar conditions, and so a calculation of the damping due to either one of them should be sufficient for present purposes. This calculation is made in Section IV, and only its more pertinent consequences will be recorded at this point.

The rate of energy dissipation is proportional to the kinematic viscosity, which in turn is almost inversely proportional to the gas density. Dissipation therefore becomes increasingly important at greater heights. Further, dissipation tends to be more important in the waves which have the smaller scale sizes. Consequently, viscosity is likely to remove from a wave train the modes with the smaller scale sizes, and, if the train is progressing upwards, to remove more and more of these modes as successively greater heights are reached.

Various criteria might be introduced at this point to establish the height at which dissipation becomes important for a given mode, but their validity would in fact depend on the particular purpose for which they were derived and on an assumption as to the region of wave generation. For the present paper, an arbitrary choice is made which should at least be satisfactory for many order-of-magnitude estimates. A viscous cutoff is assumed to occur in

those modes whose rate of energy dissipation, if maintained over a full period of the oscillation, would be just sufficient to exhaust the excess energy available in the wave motion; less dissipative modes are considered to be 'permitted', and more dissipative modes to be 'excluded'.

On this basis, it is possible to put a lower limit on the vertical wavelengths permitted in the upper atmosphere, and this limit is depicted in Fig. 3. It corresponds, incidentally, to modes of oscillation whose periods are about 10

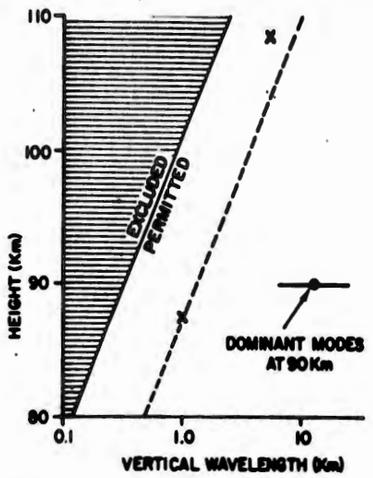

FIG. 3. Abstract of viscous effects. The 'excluded–permitted' demarkation line represents the smallest vertical wavelengths that can be propagated at meteor heights, without an overriding dissipation of energy. The demarkation line for periods of 200 minutes is shown as a broken line. The X's indicate the smallest wavelengths apparent in Fig. 1. The dominant modes of oscillation at 90 km are depicted, with a spread or possible error of a factor of 2 shown by the horizontal bar.

minutes. Modes having other periods, both greater and smaller, have somewhat greater lower limits on their permitted scale sizes, and the limits for periods of 200 minutes are illustrated in Fig. 3. (See Section IV, Fig. 11, for a more complete representation.) Also shown in Fig. 3 are the smallest vertical wavelengths found at 87- and 108-km heights, as listed in property (vi) above. It is evident that this property is readily explained as a consequence of viscosity, at least insofar as the general trend and orders of magnitude are concerned.

The dominant mode at 90-km heights is also indicated in Fig. 3, and may be seen to lie well in the permitted region; in fact, at the 90-km level, it dissipates only 1/80th of its energy each cycle. It would be rapidly damped at 110 km, however, and so must there give way to some less dissipative larger scale mode of oscillation. Herein lies a partial explanation for the general tendency towards an increase in the scale of the dominant mode at greater heights, recorded in property (vi). A full explanation depends in part of course on the spectrum of generated waves, and so can come only when the excitation process is understood.

The conflicting consequences of the exponential growth of velocity amplitude, in the absence of dissipation, and of the increasing importance of dissipative

processes at greater heights, can be resolved to some extent by suitable analysis. As an illustration, the wind system depicted in Fig. 1 has been rescaled by the following procedure: an arbitrary shearing wind has been removed, varying linearly from +20 m/s at 82 km to −42 m/s at 113 km, the residual winds have been multiplied by the square root of the atmospheric density (as derived from the ARDC model atmosphere of Minzner, Champion, and Pond (1959)) to remove the amplification factor, and the result has been plotted in Fig. 4. This figure can now be thought of as a representation of a wave train propagating from bottom to top, subject to the dissipation of its smaller scale content as it progresses, and with its unfamiliar amplitude growth suppressed.

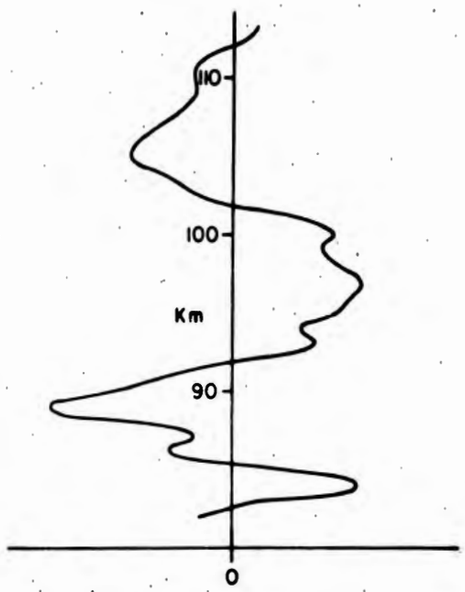

FIG. 4. Normalized wind profile at meteor heights, measured to the right and to the left from the '0' position, deduced from Fig. 1 by the removal of the general shear and the reduction of the residual by a factor proportional to $\rho_0^{\frac{1}{2}}$.

Insofar as the amplitude of the larger scale components at the top is equal to that at the bottom, this analysis confirms quantitatively the theoretical prediction of the exponential growth with height. Further, it now permits a more appropriate assessment of a representative vertical scale of the irregular winds, for the general shear has been removed and the greatly amplified oscillations at high levels no longer dominate over equally energetic modes found at lower levels. The autocorrelation coefficient of the curve in Fig. 4 falls to zero at a displacement of 4.0 km, in contrast to the displacement of 6.6 km appropriate to Fig. 1 (Liller and Whipple 1954). Even this new value must be considered to be an average over the range of heights depicted, and so most appropriate to the 97-km level. A somewhat smaller value, perhaps 3 km, should then apply at 90 km. On this reasoning, the revised autocorrelation analysis confirms the 12-km vertical wavelength quoted in property (i) above.

This completes the comparison of the theory with the meteor observations. It will be apparent already that internal atmospheric gravity waves provide an adequate empirical interpretation of the motions observed. Their further substantiation must be based on additional comparisons with observation, and on the identification of a plausible mechanism for their generation. The latter aspect will now be considered.

11.3. Wave Generation

There appear to be two major mechanisms available for the production of the waves discussed above. One of these derives its energy from the tidal oscillations of the atmosphere, and it produces the small-scale waves more or less *in situ*. The second has its origins in the wind systems of the tropospheric (or stratospheric) regions, and depends on an upward propagation of energy to meteor heights. This mechanism is in some respect preferable to that of the tides, but neither necessarily excludes the other; each will be discussed in turn.

Tidal oscillations of the upper atmosphere are now understood in broad outline and in many of their detailed features. (See Hines (1959*b*) for a short review of the subject and for pertinent references.) One of their most notable features consists of the upward amplification of the wind systems they generate, roughly in accordance with the factor $\exp(\gamma gz/2C^2)$ already noted for internal gravity waves. This amplification is arrested, in the case of the dominant semidiurnal component, in the height range 50–80 km above the earth, where most of the tidal energy is reflected downwards as a consequence of the local thermal decline. An appreciable part of the energy does penetrate to a freely propagating region above, however, and there the amplification of the winds continues as the height increases. At a height of 90 km, a representative tidal wind speed would be 20 m/s and a representative rate of growth 1.6 m/s km (or 8% per km), as revealed by meteor data.

Meteor observations indicate a change of phase with height as well, in agreement with theory. This may be extrapolated upwards to heights of 110–115 km, where it comes into agreement with the tidal phase revealed by *E*-region 'drift' observations. The tidal winds revealed by these 'drifts' are, however, still only of the order 20–30 m/s, and so are quite inconsistent with an empirical extrapolation of the meteor measurements and with the theoretical growth (see Fig. 5). The tides appear then to lose much of their energy somewhere between heights of 90 and 115 km, and this loss cannot be explained adequately by molecular transport phenomena nor by hydromagnetic damping.

The occurrence of an abnormal effect should not be surprising, however, for the nature of the tidal oscillation undergoes a severe change at the heights concerned. Whereas linear (perturbation) theory is adequate to describe the motion at lower levels, the amplification with height renders such a simplification inaccurate at 90–100 km altitudes. Density variations there are of the order of 10%, for example, and other contributions to non-linear effects are of a similar order. These values would be increased to 50–100% at the 115-km level, if the exponential growth appropriate to the linear theory should persist, so it is quite clear that new considerations must become pertinent.

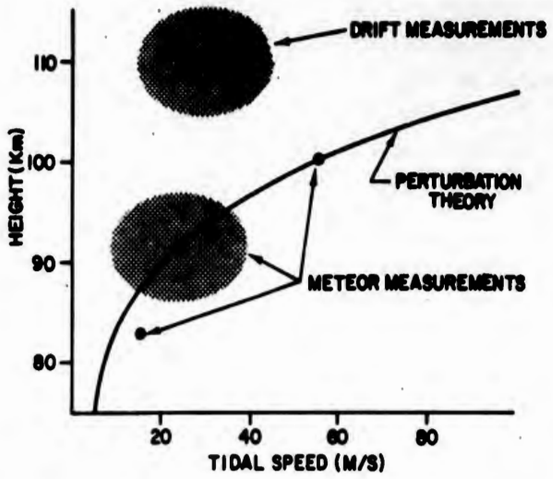

FIG. 5. Variation of tidal amplitude with height. The meteor measurements derive largely from Greenhow and Neufeld (1955), the drift measurements from Briggs and Spencer (1954). The theoretical curve is drawn on the assumption that the amplitude varies as $\rho_0^{-\frac{1}{2}}$, with ρ_0 taken from Minzner, Champion, and Pond (1959).

It seems not unreasonable to suppose that there would be a tendency towards the generation of some form of higher harmonic, or even of a broad spectrum, of related atmospheric waves. A cascade process analogous to that pictured in turbulence theory might well develop, whereby energy is passed into the smaller scale modes of propagation where it can be more readily dissipated. All modes of atmospheric oscillation might be involved in the process, but the internal gravity waves should receive a substantial portion of the cascading energy since they are of the same family as the parent tidal motions. This, then, provides one possible mechanism for the generation of the waves previously discussed.

It may not, by itself, be an adequate mechanism for the complete spectrum, however, for it appears to suffer from one limitation. The cascade process would tend to work more effectively at the greater heights, and the waves detected at (say) 90 km might then be expected to have propagated downwards from a region of generation above. But at the 90-km level the spectrum contains waves of smaller scale than can be supported at (say) the 100-km level, and they must then have been generated locally by waves of intermediate scale propagating their energy downwards from the primary interaction region. The latter waves would, however, be diminishing in amplitude as they descended, and the efficiency of the cascade process would then be falling; the likelihood of small-scale generation would be diminishing, and it is questionable whether small-scale winds with speeds comparable to those of the intermediate scales would be produced. The possible importance of this objection is as yet difficult to assess.

The second proposed mechanism is much more direct in concept. Disturbances associated with wind and weather systems in the lower atmosphere, or with possible instabilities in the middle atmosphere, may be expected to generate at their boundaries atmospheric waves which would propagate away

through the undisturbed regions. This fact has apparently not been studied, and consequently little is known as to the order of magnitude of the wave amplitudes which might arise immediately adjacent to the source regions. However, in view of the decrease of density between the ground and the 90-km level, by a factor of 5×10^5, it is evident that the amplitudes of plane internal gravity waves could increase over the same height range by a factor of 700. Accordingly the observed upper atmospheric winds could be produced by a process of wave generation in the lower atmosphere, whose associated oscillatory motions there would be only a few centimeters per second and whose fractional density variations would be about one part in 10^3 or 10^4. It does not seem unlikely that wave amplitudes of this order would in fact be produced in the troposphere quite commonly. Alternatively, waves generated in the middle atmosphere as a consequence of some instability process might provide the requisite energy, although their initial amplitudes would have to be correspondingly greater.

Propagation of energy upwards from low levels entails a further process which will be treated only briefly in the present paper: reflection of waves at a 'thermal barrier'. Temperature variations in the atmosphere modify the simple picture of wave propagation so far employed, and some of their effects are indicated formally in Section IV. It is sufficient for immediate purposes to note that some modes of oscillation (including the semidiurnal tidal component) are strongly reflected in the middle atmosphere and cannot propagate through to the upper regions except in a severely attenuated form. It is reasonable then to expect these modes to be virtually absent from the spectrum observed at 90 km, if the upward propagation of wave energy is indeed the pertinent mechanism. Since other modes of propagation are removed by viscous dissipation as previously discussed, the available spectrum could be in the end severely restricted.

These considerations form the basis of Fig. 6, wherein the available spectrum appropriate to 90-km heights is illustrated. The various modes of propagation

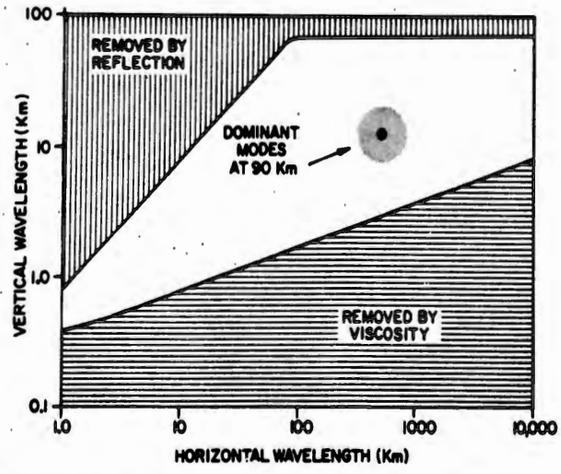

FIG. 6. The spectrum of wavelengths available at 90 km, following propagation from the lower atmosphere, is shown by the unhatched area. The dominant modes at 90 km are also depicted, with a spread or possible error of a factor of 2 indicated by the shaded region.

are here characterized by their horizontal and vertical wavelengths, although any point in the $\lambda_x - \lambda_z$ domain is of course associated with a specific period of oscillation as determined by the dispersion equation. (See Section IV, Fig. 11, for a more complete description.) The two large shaded areas in Fig. 6 represent the reflected and dissipated modes, and so the unshaded area depicts the available spectrum. The dominant observed modes are also illustrated, and are found to lie centrally in the permitted range. Clearly, low-altitude generation of the meteor-height motions must be seriously contemplated.

III. ADDITIONAL INDICATIONS OF ATMOSPHERIC WAVES

III.1. Radio-wave Scattering in the D and E Regions

Irregularities of electron density in the *D* and lower *E* regions have been studied by the scattering of radio waves at various frequencies. A good deal of agreement exists between different workers as to the most important scattering heights in the normal ionosphere. In an idealized picture, these might be taken to lie immediately above a base in the 80–95 km range and immediately above a second base in the 65–75 km range (Gregory 1956). For the moment, attention will be confined to the latter.

Observations of the lower scattering region obtained at low frequencies might best be interpreted in terms of a ledge of ionization at the appropriate height, due, perhaps, to a quasi layer formation. At medium frequencies, however, the evidence suggests that some further process must be operative. The fading behavior and range distribution of the returned signals are such as to suggest that the scattering is produced by a few irregularities at a time, and that these irregularities are appreciably anisotropic (Gardner and Pawsey 1953; Fejer and Vice 1959). Dieminger (1955) has observed day-to-day variations in the occurrence of the echoes, and this has led him to suggest that some form of atmospheric meteorology is likely to be involved in the process quite apart from any regular layer formation that might be envisaged. It is tempting to look for an interpretation of the anisotropy and of the meteorology in terms of internal atmospheric gravity waves.

No attempt will be made here to develop such an interpretation in detail, but certain pertinent factors do bear consideration. Application of conventional scatter theory (Booker 1959) to the case of backscatter leads to the following conclusion: the important irregularities of electron distribution are those whose 'corrugation wavelengths' in a spacial Fourier analysis are equal to half the radio wavelength, or typically 75 m in the medium-frequency observations. Atmospheric wavelengths of this same magnitude then warrant initial consideration.

Reference to Section IV (Fig. 11) will reveal that, in the model atmosphere there adopted, atmospheric waves with wavelengths of 75 m cannot propagate without severe dissipation at heights above 80 km, and that, even at lower heights, dissipation is extreme except for waves propagating at an angle sufficiently off the vertical. For example, at 60 km the waves must have an off-vertical angle $> 5°$, and at 70 km this is raised to 15°; waves propagating at angles closer to the vertical are simply removed by viscous damping.

Insofar as the atmospheric waves are of importance to the medium-frequency results, then, it may be expected that the region of most intense scattering will be determined by the competing effects of dissipation on the one hand and the upward growth of electron densities and wave amplitudes ($\propto \exp[\gamma g z/2C^2]$) on the other. Indeed, the very occurrence of significant scattering may depend on the occurrence of a region of overlap between appreciable 75-m atmospheric waves and appreciable electron densities. Day-to-day variations in the atmospheric wave spectrum and perhaps in the dissipative cutoff could then provide the meteorology suggested by Dieminger, while inherent limitations on the possible region of overlap would result in the normal confinement of the scatterers to a relatively restricted range of heights. The anisotropy reported by Gardner and Pawsey and by Fejer and Vice would, on the present interpretation, imply a limited angular spectrum (or what is equivalent for these waves, frequency spectrum), since it was based on a limited spread of observed slant ranges.

The model just proposed implies two features that may be open to experimental check. Since the important atmospheric waves propagate at an off-vertical angle, the important signal returns would also come from off-vertical angles. Whether this is a measurable effect or not will depend in part on the angular spectrum of the atmospheric waves and on the relative heights of appreciable ionization and atmospheric dissipation, in addition to ordinary equipment limitation. It is further subject to some uncertainty in that the corrugation wavelengths of importance are those in the electron distribution, whereas the discussion so far has concerned only the real wavelengths of the atmospheric oscillation. The conversion is by no means direct, because of the superimposed vertical variations of wave amplitude and electron concentration. The second predicted feature is subject to fewer uncertainties: longer radio wavelengths should tend to scatter from somewhat greater heights, because the related atmospheric waves will reach maximum amplitude at greater heights, provided only that the electron concentration does not diminish appreciably.

The scattering centers observed at heights above 80 km will now be discussed. They are neither anisotropic nor severely confined in height. Moreover, their corrugation wavelengths as currently interpreted (Booker 1959; Bowles 1959) are much smaller than could be supported by propagating waves at the heights concerned. There seems to be no reason to depart from the prevalent interpretation of them as manifestations of turbulence. The turbulence is likely to arise from the winds revealed by meteors, as proposed by Booker (1956), but with a smaller input power than he suggested (Hines 1959*c*; Stewart 1959; Greenhow 1959). At the lower power, the small-scale turbulence cutoff deduced by Greenhow (1959) is compatible with the scales inferred from radio-wave scattering (Bowles 1959). A consistent picture seems now to have emerged, and in it gravity waves have no part to play except for the provision of suitable wind shears and hence of energy. In this respect, their possible variation due to meteorological conditions can still determine the intensity and extent of the turbulence, but otherwise their pertinence to the higher

small-scale scattering centers can be ignored. (The reverse may not be true, however. Turbulence of sufficient strength would introduce a significant eddy viscosity, which could dissipate the wave energy more rapidly than would molecular viscosity alone. The increase of effective viscosity is unlikely to exceed an order of magnitude at 90-km heights, and should be negligible above 100 km.)

Scattering of v.h.f. waves from the lower levels (Bowles 1959) is almost certainly an indication of turbulence there, and this turbulence might again be generated from internal gravity waves. Such a source could be questioned on the grounds that the velocities produced by these waves would be greatly reduced at the lower heights, but two counterarguments seem to be at least as pertinent. The wave spectrum would be broader at the lower levels, and would contain in particular much smaller scale waves, so the actual velocity shears could remain quite high. Moreover, because of the thermal decline at the heights in question, the atmosphere there is far less stable than at the 90–100 km levels and much smaller shears could lead to a comparable degree of turbulence. Both at the lower and at the higher levels, the production of turbulent motions will require further examination.

Scattering or partial reflection of radio waves from sheets of 'sporadic-*E*' ionization occurs irregularly, sometimes during disturbed and sometimes during quiet conditions. Various types of sporadic *E* are normally distinguished, and they may in many cases have their origin in different phenomena. Some occur at sufficiently great heights ($\gtrsim 115$ km) that the geomagnetic field may play an important role in any movement of ions under the influence of winds, to the extent that strong concentrations of ionization may be produced purely by a shear flow of the neutral gas (Dungey 1956, 1959). If this process is dominant in the production of a sporadic-*E* layer, it matters little whether the shear flow is produced by atmospheric waves or by other means; a wave theory would have little to add to an empirical description of the event. Rapid recombination would tend to lessen the effectiveness of Dungey's process, however, as would occurrence at a lower height where the geomagnetic field has less effect. The atmospheric density distribution would then become more important in determining the intensity of the ionization irregularities, and the fluctuations in it can be predicted by a wave theory. The fractional variation of density that occurs in the asymptotic approximation previously employed is given by $(\gamma-1)^{\frac{1}{2}}U_z/C$, and so is typically 5–20% at *E*-region heights. This seems to be appreciable even under normal circumstances, and it would not require an excessive increase in U_z to become quite marked. The density fluctuations associated with the observed wind shears should therefore not be overlooked in any search for mechanisms to explain sporadic *E*.

III.2. Ionization 'Drifts' in the E and F Regions

Extensive studies have been made of moving irregularities of electron distribution in the *E* and lower *F* regions. The studies are conducted by means of fixed-frequency ionospheric sounders employing spaced receivers. The ionospheric irregularities impose on the radio waves phase (and perhaps

amplitude) fluctuations of sufficient intensity to produce a strongly varying diffraction pattern in the returned power, and this pattern is sampled at three or more closely spaced sites. Analysis of the received signals reveals rapid variations of the pattern, but often a horizontal drift of the whole pattern can also be distinguished. The drift exhibits apparently random velocities, superimposed upon more constant components. The latter include tidal motions, as previously noted, whose phase is in agreement with an upward extrapolation of the phase revealed by meteors, and whose amplitude lies typically in the range 20–40 m/s. The random motions are of much the same order. (See Briggs and Spencer (1954) for an early comprehensive summary, or Ratcliffe (1959*a*) and Hines (1959*b*) for more recent outline reviews and pertinent references.)

The term 'drift' was originally applied to these observations in order to avoid any prejudgment of the mechanism underlying them. Much thought has, however, been given to the possibility that the motion is in fact a consequence of atmospheric winds blowing on the charged constituents, or alternatively a more complicated effect produced by electric fields. Both of these mechanisms rely on the prior existence of irregularities in the distribution of ionization, produced perhaps by a postulated turbulence, and they are responsible only for the motion and deformation of the irregularities. Because of the complicating effects of the geomagnetic field, the motions observed are likely to bear no simple relation to the driving wind or electric field. For this reason, it is difficult to understand how the tidal component should be revealed so clearly, evidently with the correct phase. (See Ratcliffe 1959*a*.)

In addition to the atmospheric winds and electric fields, progressing waves are recognized as a possible cause of the drift motions. This interpretation has many appealing features, particularly now in view of the meteor evidence for the existence of internal atmospheric gravity waves.

If atmospheric waves are the causal agency, they produce their own irregularities as they travel and do not rely on the prior existence of irregularities. In this respect, it is necessary to distinguish clearly between the speeds associated with the amplitude of the oscillation (the U_x and U_z involved in distortions of meteor trails) and the phase velocity of the propagating wave. The former would indeed move any pre-existent local irregularities as if by a wind, but the latter gives the velocity of the irregularities generated by the wave itself. It is instructive to examine the possibility that the main random drift motions result basically from phase progressions, and not from the associated oscillatory movements.

This interpretation has the fortunate advantage that the complications resulting from geomagnetic influences are of little concern. Such complications may result in a pattern of electron distribution somewhat different from that of the atmospheric density, but the velocity of the two distributions would be essentially the same. The problem of interpretation is thereby considerably eased. The occurrence of the appropriate tidal component is readily explained, for the atmospheric waves might be expected to propagate more or less randomly as measured in a co-ordinate system moving with the bulk of the

gas; once the random phase motions are removed by analysis, the bulk motion would remain.

It should be noted in passing that phase velocities can be altered readily, simply by the introduction of a new wave train. Rapid reversals of velocity are often observed, and might be interpreted most easily on this basis.

Some question exists as to the height at which the irregularities impose their effects on the probing radio waves. Theoretical arguments can be advanced to suggest that the important irregularities are those that occur at the height of reflection (Booker 1955), but they are by no means compelling (Pitteway 1958). More reliable is the observational evidence that records made at two closely spaced frequencies reveal different tidal components, and that the relative phases of these are such as would be expected for the different heights of reflection. It may be assumed as a consequence of this empirical result that attention should indeed be confined to irregularities at the reflection height only (Jones 1958). On this basis the random drift speeds would be the horizontal phase speeds of the important irregularities, as illustrated in Fig. 7.

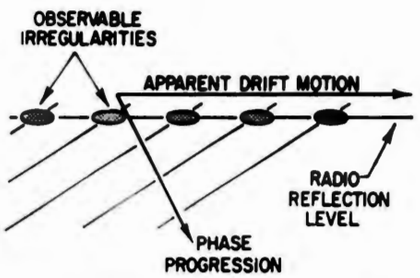

FIG. 7. Schematic representation of the apparent horizontal motion of 'drift' irregularities in the E and lower F regions, resulting from the intersection of moving phase surfaces and the stationary unperturbed level of radio reflection.

It should now be noted that the dominant modes of atmospheric oscillation are not necessarily the modes of greatest importance in the drift measurements. In view of the observational evidence just discussed, it seems reasonable to view the diffraction process in a simplified way, as a consequence of interfering signals reflected specularly from various points in a distorted surface of constant electron density. Horizontal gradients of the electron distribution then become pertinent, provided they persist over distances at least comparable to the principal Fresnel zone dimensions ($\simeq$ 2–3 km), but these gradients are unlikely to be large in the case of the dominant atmospheric waves (whose λ_x = 490 km or so). The horizontal scale of the important irregularities cannot be deduced from the present theory until more is known of the available atmospheric wave spectrum, but the observational evidence appears to support values of 10 km or so (Ratcliffe 1959*b*).

Reference to Section IV (Fig. 11) will show that, at a height of 110 km, internal atmospheric gravity waves having $\lambda_x \simeq 10$ km are confined to periods less than 18 minutes (because of viscosity) and greater than 8 or 5 minutes

(depending on the location of the energy source, below or above the thermal barrier). Their horizontal phase speeds are then confined to values greater than 9 m/s and less than 21–33 m/s. For $\lambda_r \simeq 20$ km, the limiting periods are 30 minutes and 8–5 minutes, and the corresponding horizontal phase speeds are 11 m/s and 42–66 m/s. It is evident that such values are compatible with the observations.

A general tendency towards greater speeds at greater heights exists, and can be explained as a further effect of viscosity, As an example, waves with $\lambda_r \simeq 10$ km would be severely damped at heights of 125 km, and the lower speeds they contribute to the observable spectrum would then be lost. Even the waves with $\lambda_r \simeq 20$ km are confined to periods shorter than 15 minutes, and so to horizontal speeds exceeding 22 m/s, at this height.

A tendency towards greater speeds at times of high magnetic activity has also been found (Chapman 1953). Possibly the simplest explanation of this would lie in an assumed shift of the frequency spectrum towards shorter periods, produced perhaps by local generation. Less direct interpretations can be advanced, based in part upon assumed temperature increases such as those anticipated by Dessler (1959) as a consequence of hydromagnetic heating.

III.3. Travelling Disturbances

Fixed-frequency studies of *F*-region movements are conducted not only by means of the diffraction technique just treated, but also by the identification of isolated events at widely separated stations. The observations in this case are interpreted in terms of horizontally moving ripples in the contours of constant electron density (Munro 1953, 1958), and speeds in the range 50–200 m s are quite typical. In the absence of further information, these properties might be interpreted immediately on the basis of internal atmospheric gravity waves as a direct extrapolation from the suggestions and comparisons presented in the preceding subsection.

In the case of these isolated events, however, a much more complete picture is available and a more critical assessment of any theoretical interpretation is therefore possible. The additional information comes from studies of the temporal and spacial variations associated with the disturbances, and of the horizontal distances over which they progress (Munro 1950, 1958; Munro and Heisler 1956*a*, *b*; Heisler 1958; Bibl and Rawer 1959). The quasi periods of an individual event are typically 20–25 minutes, with a lower limit of 10–12 minutes and an upper limit measured in hours. The vertical variations are such that these disturbances appear almost always (or always) to have a downward component of motion, although it has been recognized from an early date that this apparent motion might be produced by an inclined front moving horizontally. The inclination of the front is variable, but a forward tilt of 45° appears to be considered representative. The horizontal structure size is typically 150–200 km, and the vertical structure size is then about the same. The mean horizontal speed is about 125 m/s, and the fluctuation of electron density is sometimes several per cent. The disturbances are often seen in widely separated regions on broad fronts, and may progress some hundreds of kilometers during lifetimes of many hours.

The earliest attempt at a theoretical explanation was that of Martyn (1950). It was based on the assumption that the disturbances were produced by internal atmospheric gravity waves of substantially the same type as those discussed here, but further assumptions were introduced which complicated the model and led to highly debatable conclusions. Specifically, the great distances over which the disturbances often progress were taken as an indication of a trapping of wave energy below some level near or above the peak of the F layer. This led in turn to the conclusion that the waves were 'cellular' in nature: that they provided continuously rotating cells of neutral gas in the F region, and that these cells were moved horizontally by a general wind motion. The cellular waves would move the ionization in such a way as to produce the required inclination of the wave front if the disturbance were moving towards the equator.

The success of this model was limited by the direction of motion required, and by the necessity for an artificially postulated region of exceedingly low γ in or above the F layer. The model itself was subsequently abandoned in preference for a mechanism based on the effects of a dominating electric field (Martyn 1955). An alternative proposal was advanced by Dungey (1955), involving hydromagnetic waves generated in the outer reaches of the terrestrial atmosphere and propagated downwards by the ionized gas. A hydromagnetic interpretation was also suggested by Akasofu (1956), but for it the source was taken to lie in a shock process in the solar corona.

Hines (1955, 1956) returned to atmospheric gravity waves as the basic mechanism, but avoided any assumption of energy trapping. It was possible then to treat these waves in their simplest form, and to interpret the observed motion primarily as a phase progression. The large amplitudes observed in the fluctuations of electron density were stressed, and it was suggested that they might result from a selective enhancement of certain preferred modes of oscillation. The possibility of such an enhancement was investigated along two distinct lines, both of which anticipated a resonant response of the ionized constituents. In the one case, the resonance was sought in the associated electromagnetic oscillation, and a corresponding response in the electron motion was deduced; in the other the resonance was sought directly in the natural oscillations of the electron gas.

The first of these developments is of doubtful relevance except in its general commentary, while the second is most certainly incorrect. The error in the latter was pointed out by D. F. Martyn (private communication). It lies in the neglect of ion–electron attraction, and was based on the application of an inappropriate relation for the magnitude of the pertinent electric fields. The error can be corrected approximately by recasting the formalism in terms of ions and seeking once again a resonant response; the electrons would follow this response. Such a resonance does appear to be possible in principle, and it would occur at about the right temporal and spacial scales. Additional objections have, however, been raised by L. R. O. Storey and J. A. Fejer (private communications), concerning the rise-time of the resonance and the magnitudes and height variations of the resonant scales. These aspects will not be discussed

in detail here, but it may be noted in passing that they concern only the resonance postulate and not the basic mechanism of atmospheric waves.

At this point it is appropriate to insert some general comments on the nature of the travelling-disturbance mechanism. First, it is possible to argue against atmospheric winds as the transport mechanism on a priori grounds, albeit somewhat inconclusively. Winds at F-region heights are incapable of moving ionization across the geomagnetic field lines with any appreciable efficiency (see, for example, Ratcliffe 1959*a*), but they do impart to the ionization their component of velocity in the direction of the field. Horizontal winds moving with the speeds observed would tend thereby to raise or lower the whole F layer by some tens of kilometers in the course of a single disturbance, and no such unidirectional motion is normally observed. This vertical motion could be diminished by the effects of vertical polarization fields generated by the motion itself, but the degree of diminution is not appreciable at most latitudes. Nearly complete three-dimensional polarization fields would have to be established to produce a sufficient change, and they are unlikely to occur on the spacial scales required. Furthermore, a most crucial observation has recently been reported: one disturbance has been found to pass right through a second, and each has maintained its independent identity and velocity despite this juxtapositioning (Heisler 1958).

The foregoing objections apply equally well to models in which the general motion is attributed to the effects of a dominating electric field in place of a dominating wind. They do not apply to a wave hypothesis, however. Waves can, of course, pass through a medium and distort it as they go, without producing any enduring displacement of it. They can also pass through one another without significant modification, provided only that their intensity is not so great that intermodulation effects become serious. Furthermore, as a more positive argument, it may be noted that the quasi periodicities are often clearly apparent and that they can be explained readily only on the basis of a wave motion.

Hydromagnetic waves of the pertinent periods are likely to be strongly absorbed in the F region, but this in itself does not eliminate them from consideration. They can probably be discounted on the basis of speed, however, for they travel much faster than the observed disturbances. The Alfven speed at 225 km is about 2000 m/s, and this would only be increased if the ion mass density were taken to be pertinent in making the calculation, in place of the neutral gas density, as would be appropriate at shorter periods. Moreover, one of the chief attractions of the hydromagnetic hypothesis is its ability to explain the apparent downward progression of the disturbances as a direct consequence of generation at great heights. Such a picture conflicts with recent observations, however, which indicate that disturbances in the $F-1$ region are often not to be found in the $F-2$ region above (Heisler 1958).

There remain for discussion the internal atmospheric gravity waves with which the present paper is concerned. These waves most certainly do have the necessary characteristics.

The shortest period of oscillation of these waves, if the atmosphere is taken

to be isothermal, is τ_g, and a representative value for this parameter at a height of 225 km would be about 15 minutes. This is close to the shortest periods observed, and it is, of course, subject to modification as the temperature and density models are changed. (In fact, somewhat shorter periods can be present in a region of temperature incline, and a more appropriate theoretical value for the cutoff period is about 12 minutes, as observed.) The longest periods are probably determined by hydromagnetic viscosity, which would rapidly dissipate the energy of waves whose periods were comparable to $2\pi\rho/\sigma_1 B_0^2$, where ρ is the atmospheric density, σ_1 is the Pederson conductivity, and B_0 is the geomagnetic induction (all in rationalized m.k.s.a. units). A representative value for this period would be about one hour during the day, although it would change appreciably through the depth of the F region and from time to time.

If the energy of the wave were propagating upwards in the F region—and this point will receive further discussion—it would ultimately be dissipated when viscous effects became severe. These might be either molecular or hydromagnetic in nature, depending on the mode of oscillation and on the relative magnitudes of various parameters. The point will not be pursued here in detail, but it obviously provides an explanation for the disappearance of disturbances between the $F-1$ and the $F-2$ regions. A complementary observation, that travelling disturbances seen in the F region are normally not also seen in the E region (Martyn 1959), is explicable as a consequence of the exponential growth of amplitude with height which characterizes internal gravity waves. The amplitude at 100 km would be only about 1/30th of that at 200 km, for example, and so would produce density variations of only a fraction of 1% at the lower level. This estimate may be contrasted with that previously calculated for other waves present at E-region heights; the travelling-disturbance waves would simply be lost in the 'noise' of the broader and more intense spectrum that is available there.

The horizontal and vertical structure sizes are, of course, related by a dispersion equation, which is derived in the next section. The asymptotic form employed previously, equation (2), is not sufficiently accurate for present purposes, but it may be generalized somewhat to

$$(\tau^2-\tau_g^2)\lambda_z^2 = \tau_g^2\lambda_x^2, \tag{4}$$

which is valid provided λ_z is not too great ($\lambda_z \lesssim 500$ km, say). This indicates that a wave front will be inclined at 45° if $\tau^2 = 2\tau_g^2$, or if $\tau \simeq 20$ minutes here. This conclusion is certainly compatible with the observations.

The possible range of horizontal and vertical structure sizes is limited by viscosity as before, and by the thermal barrier if the energy source is taken to reside in the lower atmosphere. The limits are indicated in detail in the next section (specifically, in Fig. 12), but they may be summarized here briefly in terms of the corresponding limits on the horizontal speed of progression. The lower limit is set by viscosity, and is about 14 m/s (0.8 km/min); the upper limit is determined by the thermal barrier, and is 170 m/s (10 km/min) on the model adopted. These conclusions may be compared with

the observational results of Munro (1958): "Very few values below 3 km/min are recorded, though there is no technical reason why they would not be observed. In all, 54 have been recorded in 7 years, ranging from 1.9 to 2.9 km/min. Values greater than 15 km/min are relatively scarce, and those recorded are considered doubtful since the possible errors become great at speeds above 10 km/min." When allowance is made for the uncertainties and normal variations in the parameters of the theory, a closer agreement could scarcely be expected.

The possibility that the wave energy progresses upwards from the lower atmosphere warrants further discussion, in view of the fact that the apparent vertical motion is normally downwards. One of the peculiarities of internal atmospheric gravity waves is that they propagate energy upwards only in modes whose phase progression is downwards, although the horizontal components of energy and phase velocity are directed in the same sense. (See next section, where also are indicated the exceptions to this statement that can arise when temperature gradients are present.) Accordingly, the postulate that the energy is provided at low levels leads immediately to the theoretical conclusion that a disturbance front in the *F* region will be tilted forwards at the top. When the disturbance progresses horizontally, this same postulate leads in turn to the conclusion that an apparently downward motion would be determined from recordings at a single location, in agreement with the observations.

The horizontal progression should in general not be interpreted as a horizontal flow of energy in the *F* region, however. Indeed, at periods substantially greater than τ_g, energy in the *F* region is flowing obliquely upwards along lines nearly parallel to the disturbance front, and so cannot carry that front horizontally. Instead, the horizontal progression must normally be attributed to the fact that, at lower levels, the energy flow is much more closely confined to quasi-horizontal directions. Energy propagating upwards from the lower atmosphere is refracted into a more horizontal course as it reaches the thermal decline, at heights of 50–80 km, and indeed it may be strongly reflected there and returned to the lower levels as has already been indicated. The energy that passes through this thermal barrier can of course proceed on upwards, but it may be carried horizontally over great distances before its escape is effected. The process is closely analogous to that arising in the oblique propagation of radio waves through the ionosphere, at an elevation slightly greater than the penetration angle of one of the layers. It is illustrated in Fig. 8 for the case of atmospheric waves that just succeed in passing through the barrier. (While the paths depicted there are in effect ray paths, and only transmitted rays are shown, a full-wave treatment of the problem is of course far more appropriate. It will not be attempted here, however.)

Many, and perhaps most, of the smaller travelling disturbances detected in the *F* region may well result from modes of propagation that are not too severely influenced by these refractive effects. But, if so, their apparent horizontal motions should be interpreted as horizontal phase progressions within a broad and relatively stationary front.

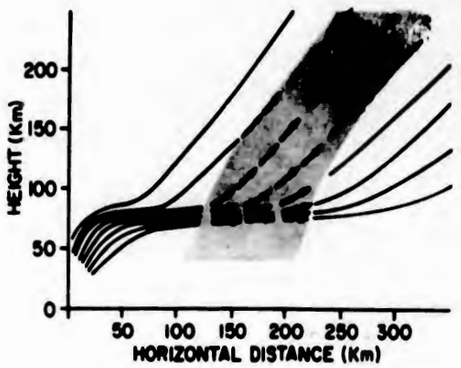

FIG. 8. Sketch of the energy flow pattern envisaged for large-scale travelling disturbances. The flow lines are shown by solid lines, and the instantaneous flow pattern by superimposed arrows. The disturbance is confined at any one time to a region such as that within which the arrows appear. It has maximum amplitude towards the top, just before dissipation becomes important, with a secondary enhancement near 80 km where the flow lines are constricted by ducting.

In those cases where a clearly defined front progresses horizontally over very great distances, on the other hand, it seems that very severe refraction must exist in the 70–80 km height range. This refraction amounts almost to a ducting of the energy, although the 'duct' is admittedly 'leaky'. This low-level ducting appears to meet the objectives originally set out by Martyn, but it does so with none of the difficulties encountered by him. The waves of present interest are those that just escape reflection at a well-known barrier in the middle atmosphere, rather than those that suffer reflection at a postulated barrier in the F region, and they entail no general wind motion nor abnormally low γ.

For simplicity, it is convenient to focus attention on the limiting modes that propagate their energy exactly horizontally in the duct. These are non-dispersive modes, which may itself be a factor of importance in their observational identification. They have a common horizontal phase speed equal to

$$(\lambda_z/\tau)_{\text{duct}} \equiv [(2C/\gamma)(\gamma-1+g^{-1}dC^2/dz)^{\frac{1}{2}}]_{\text{min}} \tag{5}$$

in which the subscript 'min' is intended to imply that the bracketted quantity is evaluated at the height where it is a minimum. (See next section, equation (55).) This is the upper limiting speed of 170 m/s previously quoted, as deduced for one particular atmospheric model, and it may be compared with the results listed by Heisler (1958) for a series of eight disturbances that progressed over very great distances. His values range from 5.8 to 12.4 km/min, or from 97 to 207 m/s, in general agreement with the theory.

An exception to the foregoing arguments should be noted. At periods very close to the local τ_g, a ducting can occur in the F region itself and the ducted waves would propagate horizontally more slowly than modes ducted by the middle atmosphere. In this case, however, the disturbance front would be nearly vertical. The possible importance of this exceptional case can only be evaluated by further observations, but it may provide an explanation for

some of the lower velocities observed, including that of the slower member of the pair reported by Heisler.

In the more general case, where the periods appreciably exceed the *F*-region τ_g, there remains the distinct and interesting possibility that *F*-region observations of ducted modes will provide direct information on conditions in the middle atmosphere.

IV. MATHEMATICAL DEVELOPMENT

IV.1. Outline of Analysis

It is the purpose of this section to assemble the mathematical relations on which many of the preceding arguments were based. A relatively simple approach to the physics of the problem is sufficient for this purpose, although such an approach must limit the ultimate applicability and accuracy of the results. Some of the complicating features of the actual process are therefore also discussed and quantitatively assessed.

The main development is based on an idealized model in which the atmosphere is taken to be stationary in the absence of waves, and to be uniform in both temperature and composition. Superimposed wave motions are assumed to have only perturbation magnitude, in that terms of the second and higher order in the oscillation amplitude are neglected in the equations governing the motion, and they are also assumed to occur adiabatically. Forces due to pressure gradients, gravity, and inertia are alone treated explicitly, and the gravitational field is taken to be constant both in direction and in magnitude.

The waves that can propagate in such circumstances are first established in outline, and two special types are discussed at greater length. One of these, termed 'internal', is then selected for further study and found to contain two distinct sequences dependent on the frequency of the oscillation. One sequence is conveniently described as 'acoustic', and it includes ordinary sound waves as a limiting class; the other is the 'gravity' sequence whose properties have been applied in the preceding sections. Some of the general characteristics of the internal waves are discussed, and formulae applicable to them in limiting cases are derived.

The effects of atmospheric viscosity are then considered briefly, and a preliminary criterion is established in order to assess their importance. The consequences of non-linear processes are also discussed, and an evaluation is made of their significance in the upper atmosphere. Finally, one complication introduced by temperature variations of the unperturbed atmosphere is described, and its relation to the arguments of the earlier sections is indicated.

IV.2. Atmospheric Oscillations in the Presence of Gravity

Under the assumptions listed immediately above, atmospheric oscillations are governed by the following set of equations:

$$\rho_0(\partial \mathbf{U}/\partial t) = \rho \mathbf{g} - \text{grad } p \tag{6}$$

$$(\partial p/\partial t) + \mathbf{U}\cdot\text{grad } p_0 = C^2[(\partial \rho/\partial t) + \mathbf{U}\cdot\text{grad } \rho_0] \tag{7}$$

and

$$(\partial \rho/\partial t) + \mathbf{U}\cdot\text{grad } \rho_0 + \rho_0 \,\text{div}\, \mathbf{U} = 0. \tag{8}$$

These are the linearized equations of motion, of adiabatic state, and of continuous mass conservation. They relate the perturbation velocity **U**, the perturbed atmospheric pressure and density p and ρ, the corresponding unperturbed values p_0 and ρ_0, the gravitational acceleration $\mathbf{g}$, and the 'speed of sound' C. The latter is related to p_0 and ρ_0 through

$$C^2 = \gamma p_0/\rho_0 \tag{9}$$

where γ is the usual ratio of specific heats; γ, C, and $\mathbf{g}$ are all constant under the earlier assumptions.

The static condition

$$0 = \rho_0\mathbf{g} - \text{grad } p_0 \tag{10}$$

may be derived from (6), and then combined with (9) to show that

$$p_0, \rho_0 \propto \exp(-z/H) \tag{11}$$

where

$$H \equiv C^2/\gamma g \tag{12}$$

is the 'scale height' of the atmosphere and z is a Cartesian co-ordinate measured in the (upward) direction opposed to $\mathbf{g}$.

Wave solutions of the set (6)–(8) may be found in complex Fourier form, such that

$$(p-p_0)/p_0P = (\rho-\rho_0)/\rho_0R = U_z/X = U_z/Z = A \exp i(\omega t - K_x x - K_z z) \tag{13}$$

where P, R, X, Z, and A are all constant, ω is a real constant, the circular frequency of the wave, K_x and K_z are corresponding constant complex wave numbers, and x is a (horizontal) Cartesian co-ordinate measured in any direction perpendicular to the z axis.

The wave numbers appearing in (13) are related to the wave frequency by the dispersion equation

$$\omega^4 - \omega^2C^2(K_x^2+K_z^2) + (\gamma-1)g^2K_x^2 + i\gamma g\omega^2K_z = 0, \tag{14}$$

which can be derived as a necessary condition for (13) to be a valid solution of the set (6)–(8). Sufficient further conditions are provided by what may be termed the 'polarization relations',

$$P = \gamma\omega^2K_z - i\gamma g\omega^2/C^2 \tag{15}$$

$$R = \omega^2K_z + i(\gamma-1)gK_x^2 - i\gamma g\omega^2/C^2 \tag{16}$$

$$X = \omega K_xK_zC^2 - ig\omega K_x \tag{17}$$

$$Z = \omega^3 - \omega K_x^2C^2, \tag{18}$$

which determine the relative magnitudes and phases of the pressure and density variations and of the horizontal and vertical components of atmospheric motion.

In the absence of gravity, (14) would reduce to the simple form

$$\omega^2 = C^2(K_x^2+K_z^2) \tag{19}$$

which governs sound propagation. For most problems normally encountered, the appropriate solutions of (19) would be those for which K_x and K_z were purely real, when the waves would indeed be simple sound waves propagating with speed C in an arbitrary direction determined by the orientation of the x axis and by the ratio K_z/K_x.

In the presence of gravity, however, equation (14) indicates that no solution exists in which both K_x and K_z are purely real and different from zero. It seems most appropriate in these circumstances still to seek a solution in which K_x is real—say $K_x = k_x$, as an explicit indication of this—and then to investigate the suitability of the resultant K_z. This choice is equivalent to an a priori assumption that the pertinent boundary conditions in any problem of interest would not introduce a requirement for an exponential decrease of amplitude in horizontal directions—an assumption which is normally made and rarely questioned in the gravity-free case if, as here, effects leading to the dissipation of energy are neglected. This same assumption underlies the neglect of a possible $K_y y$ term in the exponential of equation (13), since such a term can be removed by a rotation of the co-ordinate system about the z axis whenever the ratio K_y/K_x is real.

Equation (14) may now be examined under the assumption that $K_x = k_x$ (real). It reveals two possibilities: either K_z is purely imaginary, or else

$$\begin{aligned} K_z &= k_z+i\gamma g/2C^2 \\ &= k_z+i/2H \end{aligned} \tag{20}$$

where k_z is purely real. The first alternative is particularly appropriate to problems involving a horizontal boundary, say in a stratified (two-fluid) medium, where horizontally propagated 'surface' waves having vertical attenuation may be expected. But, insofar as it permits no variation of phase with height, it seems singularly inappropriate for application to the problem of meteor-train deformations or to that of large-scale travelling disturbances. Much more appropriate for these applications is the second alternative, which does permit a vertical component of phase variation and so of propagation. This essential feature will be emphasized here by characterizing the associated waves as 'internal', and only these waves will receive further consideration.

The family of internal waves is familiar in acoustic studies, for it includes sound waves in a generalized form imposed by the effects of gravity. It is not wholly unfamiliar at lower frequencies, for some attention has been directed towards it in relation to atmospheric motions, particularly of the troposphere (Martyn 1950; Gossard and Munk 1954, for example), and the well-known atmospheric tides in fact constitute a low-frequency limiting member (albeit subject to additional complications). Nevertheless, a more widespread acquaintance with these waves and their properties seems now to be desirable, and the remainder of this section is intended as a first step towards that end (see also Eckart 1960).

IV.3. General Characteristics of Internal Waves

In the analysis of internal waves, it is convenient to extract the common term $i/2H$ from the complex vertical wave number of all modes, and to deal henceforth only with the real wave numbers k_x and k_z. By substitution of (20) into (14), the dispersion relation can then be written as

$$\omega^4-\omega^2C^2(k_x^2+k_z^2)+(\gamma-1)g^2k_x^2-\gamma^2g^2\omega^2/4C^2 = 0, \tag{21}$$

where the solution to the basic equations is now taken to have the form

$$\begin{aligned}(p-p_0)/p_0P &= (\rho-\rho_0)/\rho_0R = U_x/X = U_z/Z \\ &= A\exp(z\gamma g/2C^2).\exp i(\omega t-k_xx-k_zz).\end{aligned} \tag{22}$$

Similarly, the polarization factors (15)–(18) may be rewritten as

$$P = \gamma\omega^2[k_z-i(1-\gamma/2)g/C^2] \tag{23}$$

$$R = \omega^2k_z+i(\gamma-1)gk_x^2-i\gamma g\omega^2/2C^2 \tag{24}$$

$$X = \omega k_xC^2[k_z-i(1-\gamma/2)g/C^2] \tag{25}$$

$$Z = \omega[\omega^2-k_x^2C^2] \tag{26}$$

in keeping with the new notation.

It is useful first to examine the dispersion relation, equation (21). This reveals that any pair of real wave numbers (k_x, k_z) can be associated with either of two distinct values of ω^2, and so with either of two values of ω if attention is confined to positive roots. Further examination will reveal that one of these values of ω is necessarily greater than

$$\omega_a \equiv \gamma g/2C \tag{27}$$

while the other is necessarily less than

$$\omega_g \equiv (\gamma-1)^{\frac{1}{2}}g/C. \tag{28}$$

It may further be confirmed that $\omega_a > \omega_g$, since γ is necessarily less than two. This indicates that two distinct sequences of internal waves can occur, the one at high frequencies ($\omega > \omega_a$) and the other at low ($\omega < \omega_g$), and that a gap in the frequency spectrum exists between them ($\omega_g < \omega < \omega_a$) in which no internal waves can be propagated. The two sequences will here be termed 'acoustic waves' and 'internal gravity waves' respectively, although these names are somewhat misleading in that pressure gradients and gravitational forces do play some part in them both.

The two sequences can be distinguished clearly in Fig. 9, which is simply a pictorial representation of the dispersion relation. In it are plotted contours of constant ω (or of constant period, τ, given by $\tau \equiv 2\pi/\omega$) in the k_x-k_z domain, using values of γ, C, and g which are representative of meteor heights. The family of ellipses in this diagram represents the sequence of acoustic waves, while the superimposed family of hyperbolae represents the internal gravity waves.

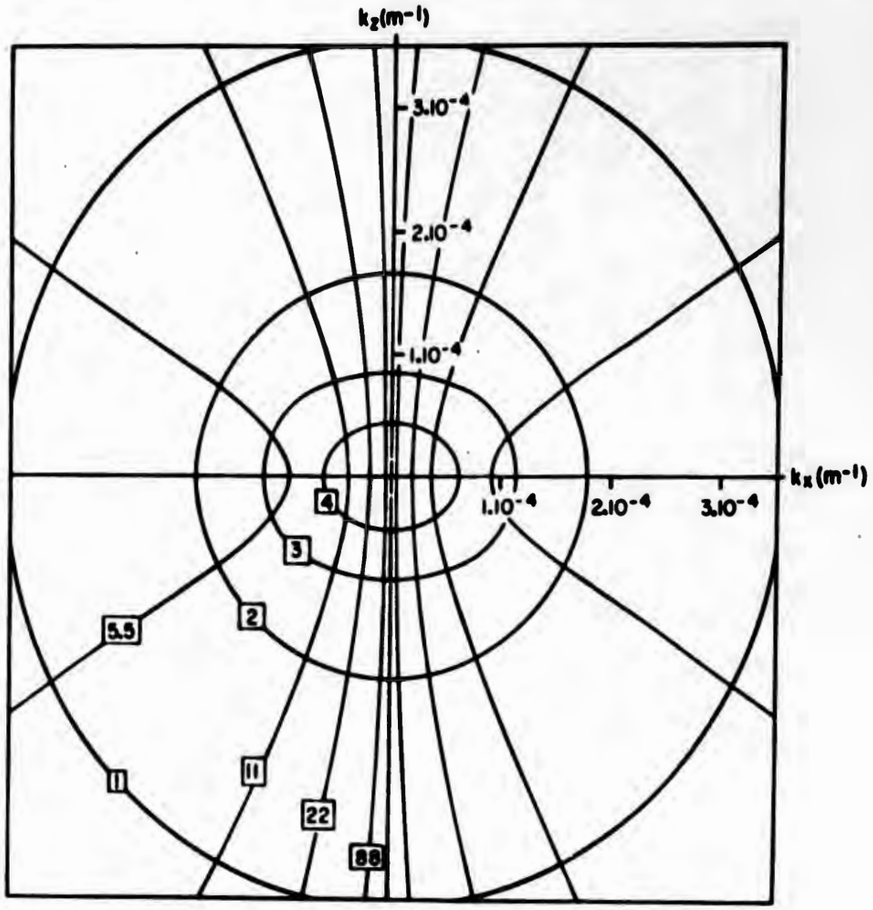

FIG. 9. Contours of constant period in the k_x-k_z domain. The periods, measured in minutes, are shown in boxes on the corresponding curves. The basic parameters adopted are $\gamma = 1.40$, $g = 9.5$ m/s^2, and $H = 6.0$ km, whence $\tau_a = 4.4$ minutes, $\tau_g = 4.9$ minutes.

The two sequences may be distinguished even more clearly in Fig. 10, where they are in fact completely separated. This diagram contains essentially the same information as does Fig. 9, but the co-ordinates are now normalized to

$$n_x \equiv k_x C/\omega, \qquad n_z \equiv k_z C/\omega \tag{29}$$

and only one quadrant of the n_x-n_z domain has been shown.

One property of the internal wave systems is made readily apparent by the representation of Fig. 10, namely, the speed of phase propagation in any given direction at any specified frequency or period. Waves propagating at an angle θ above the horizontal, for example, have phase variations in the θ direction which are governed by the wave number

$$\begin{aligned} k &= k_x \cos\theta + k_z \sin\theta \\ &= (k_x^2 + k_z^2)^{\frac{1}{2}}, \end{aligned} \tag{30}$$

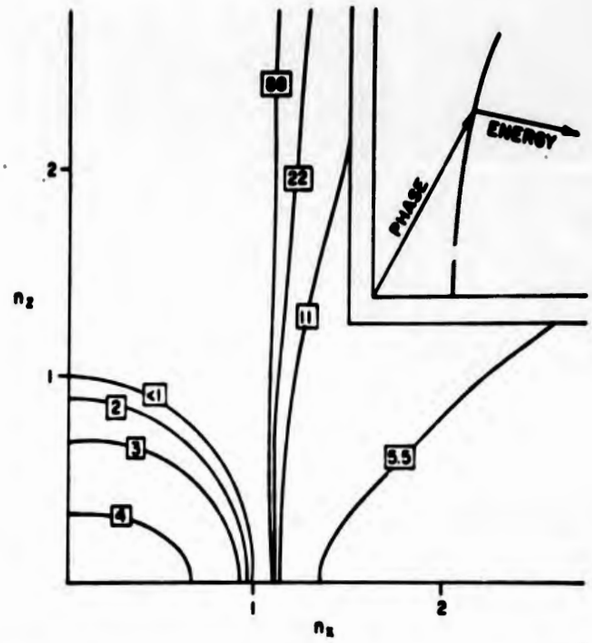

FIG. 10. Contours of constant period in the $n_x - n_z$ domain. The periods, measured in minutes, are shown in boxes on the corresponding curves. The basic parameters are the same as for Fig. 9. The relation between phase and energy progression is indicated by the geometrical construction depicted in the inset diagram.

while the speed of phase propagation in this direction is

$$\omega/k = \omega[k_x^2+k_z^2]^{-\frac{1}{2}} = C/n \tag{31}$$

say, where

$$n \equiv [n_x^2+n_z^2]^{\frac{1}{2}} \tag{32}$$

is the fluid-dynamical equivalent of the optical 'refractive index'. This refractive index can be measured as the radial distance from the origin to the appropriate τ-curve, in the θ direction. The circle in Fig. 10, at $n^2 \equiv n_x^2+n_z^2 = 1$, then represents waves that propagate with the classical speed of sound C, while contours lying closer to the origin ($n < 1$) depict waves that travel faster than C and those lying farther from the origin ($n > 1$) depict waves that travel slower than C. It is evident that acoustic waves may propagate much more rapidly at frequencies close to ω_a than they do in the more conventional high-frequency limit, whereas internal gravity waves in general propagate much more slowly.

Figure 10 also permits the ready determination of the direction of energy flow. This is given by the direction of the normal to the corresponding τ-contour, drawn at the point on that curve appropriate to a selected direction of phase propagation. The geometrical construction required is illustrated by the inset diagram, and it is based on relations derived from the concept of wave packets. These packets are formed by the superpositioning of modes

having a range of $k_x - k_z$ values about the chosen central values, and they provide a confined distribution of energy density whose maximum moves with the velocity components $[\partial\omega/\partial k_x]_{k_z}$ and $[\partial\omega/\partial k_z]_{k_x}$ in the x and z directions respectively. The slope of the velocity vector is then $[\partial\omega/\partial k_z]_{k_x}/[\partial\omega/\partial k_x]_{k_z}$. This equals $-[\partial k_x/\partial k_z]_\omega$ identically, or $-[\partial n_x/\partial n_z]_\omega$, which is the slope of the normal to the τ-contour as indicated above. The sense of the flow along this normal can be determined from either $[\partial\omega/\partial k_x]_{k_z}$ or $[\partial\omega/\partial k_z]_{k_x}$ independently, and in the present case it will be found that the appropriate normal is that drawn on the side of the τ-contour away from the origin as shown. The same conclusion as to direction and sense of energy flow may be derived by averaging the real energy flux vector, $p\mathbf{U}$, over a cycle.

It will be seen that the energy of the acoustic waves propagates in much the same direction as does the corresponding phase pattern, although there is some tendency towards a more vertical flow. On the other hand, the energy of internal gravity waves may propagate in directions radically different from the corresponding phase normals, by as much as 90° when the asymptotes of the τ-contours are approached, and for the longer periods the flow tends to be nearly horizontal. It should be noted that, in internal gravity waves, the vertical component of energy flow is opposite in sense to the vertical component of phase progression; an upward transfer of energy is accomplished by modes whose phases propagate downwards, and vice versa.

The dispersion relation (21) can be rewritten conveniently in terms of ω_a and ω_g, as

$$(\omega^2-\omega_a^2)\omega^2/C^2-\omega^2(k_x^2+k_z^2)+\omega_g^2k_x^2 = 0. \tag{21'}$$

The high-frequency limit of this equation (valid for $\omega^2 \gg \omega_a^2$) is, of course, the usual relation for sound propagation:

$$\omega^2 = C^2(k_x^2+k_z^2).$$

The corresponding waves are not freed from all effects of gravity, however, for the exponential amplitude factor still applies and the polarization ratios differ from their usual form whenever $k_z \lesssim 1/H$. These conclusions are well known in the study of sound refraction in the lower atmosphere.

Simple relations may be obtained for internal gravity waves under the asymptotic conditions which arise for $k_z^2 \gg \omega_a^2/C^2$. The approximate dispersion relation

$$\omega^2k_z^2 \simeq (\omega_g^2-\omega^2)k_x^2 \tag{33}$$

is an example. Most of these relations become even simpler at low frequencies:

$$\omega^2k_z^2 \simeq \omega_g^2k_x^2 \tag{34}$$

$$Z/X \simeq -k_x/k_z \tag{35}$$

and

$$R/X \simeq i(\gamma-1)^{\frac{1}{2}}C^{-1} \tag{36}$$

whenever $k_z^2 \gg \omega_a^2/C^2$ and $\omega \ll g/C$. All of these results have been employed already in somewhat modified forms: (33) in (4), (34) in (2), (35) in (3), and (36) to provide a measure of $(\rho-\rho_0)/\rho_0 U_z$. It is interesting to note that (35) implies an atmospheric oscillation transverse to the direction of phase propagation, albeit parallel to the direction of energy flow.

IV.4. Dissipation in Internal Waves

Although dissipative effects have so far been neglected in this section, their consequences are important in application as has been seen. The effects of molecular viscosity can be treated by the introduction of an appropriate force density in equation (6), while thermal conduction can be taken into account only by a modification of equation (7) and by the addition of a heat-flow equation. Such extensions introduce additional analytical complications, and they lie beyond the scope of the present paper.

In order to provide a preliminary assessment of the importance of these dissipative effects, however, the following procedure is adopted. The mean rate of energy dissipation through viscous losses is calculated, on the assumption that it is small, and it is compared with the mean excess energy available in the wave motion; the comparison is summarized by noting those modes whose energy loss per cycle, as calculated in this way, equals the energy available. This procedure neglects losses due to thermal conduction, but they are normally comparable in magnitude to those due to viscosity. It also neglects any modification produced by dissipation in the dispersion and polarization relations already derived, and the comparison is therefore not quantitatively correct in modes as dissipative as those quoted in summary. The latter should nevertheless represent roughly the modes whose amplitude decreases by a substantial fraction within a wavelength, and they certainly give an extreme limit to the set of modes in which dissipation can be neglected. Other criteria will be more pertinent in studies more detailed and more specific than those of the present survey.

The rate of dissipation of energy per unit mass, through the action of viscosity, is given quite generally (Lamb 1945) by

$$R = \eta(2a^2+2b^2+2c^2+d^2+e^2+f^2)-(2\eta/3)(a+b+c)^2 \tag{37}$$

where

$$\begin{aligned} a &= \partial U_x/\partial x, \qquad b = \partial U_y/\partial y, \qquad c = \partial U_z/\partial z \\ d &= \partial U_z/\partial y+\partial U_y/\partial z \\ e &= \partial U_x/\partial z+\partial U_z/\partial x \\ f &= \partial U_y/\partial x+\partial U_x/\partial y. \end{aligned} \tag{38}$$

In these expressions, η is the kinematic viscosity while $\mathbf{U}$ is the real velocity vector and not the complex representation thereof. The latter can be usefully employed, however, in calculating the mean value of R, say $\bar{R}$:

$$\bar{R} = \eta(\tfrac{1}{2})[2aa^*+2bb^*+2cc^*+dd^*+ee^*+ff^*-(\tfrac{2}{3})(a+b+c)(a+b+c)^*] \tag{39}$$

where * indicates the complex conjugate, and $a, b, \ldots f$ are given again by (38) but with complex **U**.

In the present case, $U_y = \partial/\partial y = 0$, while U_z and U_x are to be taken as given by (22) with U_z/U_x given by (25)/(26). These facts may be introduced into (39) to derive

$$2k_x^2[k_x^2C^4+(1-\gamma/2)^2g^2]\bar{R}/\eta U_zU_z^* = [k_x^2C^4+(1-\gamma/2)^2g^2][4k_x^4/3 + k_x^2k_z^2+\gamma^2g^2k_x^2/4C^4]+[\omega^2-k_x^2C^2][2k_x^2k_z^2C^2/3+5\gamma(1-\gamma/2)g^2k_x^2/3C^2] +[\omega^2-k_x^2C^2]^2[k_z^2+4k_x^2/3+\gamma^2g^2/3C^4]. \tag{40}$$

This can be reduced in the case of the high-frequency sound waves ($\omega^2 \gg \omega_a^2$) to

$$\bar{R} \simeq (2\eta/3)k^2\mathbf{U}.\mathbf{U}^* \tag{41}$$

or, for the lower frequency internal gravity waves, to

$$\bar{R} \simeq (\eta/2)(k_x^2+k_z^2)^2k_z^{-2}U_zU_z^* \tag{42}$$

in the asymptotic limit ($k_z^2 \gg \omega_a^2/C^2$).

The energy content per unit mass of atmosphere is given quite generally (Lamb 1945) by

$$\mathfrak{E} = [gz+\mathbf{U}^2/2+(\gamma-1)^{-1}p/\rho] \tag{43}$$

as measured with respect to a stationary element of atmosphere at an arbitrary height ($z = 0$). The excess due to the wave perturbation is, then,

$$\mathfrak{E}' = gz'+\mathbf{U}^2/2+(\gamma-1)^{-1}[p/\rho-p_0/\rho_0] \tag{44}$$

where primed quantities indicate changes from the rest state of the atmosphere. The mean excess is found by averaging $\mathfrak{E}'$ over a complete cycle, following the motion of the gas. While (42)–(44) involve the real perturbations, the complex representations thereof may again be usefully employed in evaluating the mean of $\mathfrak{E}'$, say $\overline{\mathfrak{E}}'$; it may be shown that

$$2\omega^2k_x^2[k_x^2C^4+(1-\gamma/2)^2g^2]\overline{\mathfrak{E}}' = [\omega^2-k_x^2C^2][\omega^4-(\gamma-1)g^2k_x^2]U_xU_x^*. \tag{45}$$

This yields the high-frequency relation

$$\overline{\mathfrak{E}}' \simeq \mathbf{U}^2/2 \tag{46}$$

appropriate to sound waves ($\omega^2 \gg \omega_a^2$), and the asymptotic relation (for $k_z^2 \gg \omega_a^2/C^2$)

$$\overline{\mathfrak{E}}' \simeq [(\gamma-1)g^2k_x^2/2\omega^2k_z^2C^2]U_zU_z^* \tag{47}$$

in the case of internal gravity waves. At low frequencies ($\omega^2 \ll \omega_g^2$), equation (47) reduces to the same form as (46).

The importance of dissipation is now to be summarized by selecting explicitly those modes in which

$$\tau\bar{R} = \overline{\mathfrak{E}}', \tag{48}$$

although it is fully recognized that the formulae already employed lose their quantitative validity in modes as dissipative as these. In application to internal gravity waves, (48) can be reduced to

$$2\pi\eta k_z^3\omega_g^2/\omega^3 = 1 \tag{49}$$

under asymptotic conditions ($k_z^2 \gg \omega_a^2/C^2$), and this is sufficiently accurate for the most pertinent calculations.

The modes so selected are displayed in Fig. 11, in what is something of a hybrid plot. The family of τ-contours already depicted in Fig. 9 is illustrated again here, although logarithmic co-ordinates are now employed and attention is confined to the $+k_x$, $-k_z$ quadrant. (The vertical co-ordinate is taken to be $-k_z$, increasing upwards, in order that the curves may apply directly to modes whose energy flows upwards; but the numerical relationships are not invalidated if this co-ordinate is read as $+k_z$.) The τ-contours are drawn for the set of parameters $\gamma = 1.4$, $g = 9.5$ m/s^2, and $H = 6.0$ km, whence $C = 280$ m/s, $\tau_a = 4.4$ minutes, and $\tau_g = 4.9$ minutes, and they therefore apply approximately to both the 70-km level and the 100-km level of a model

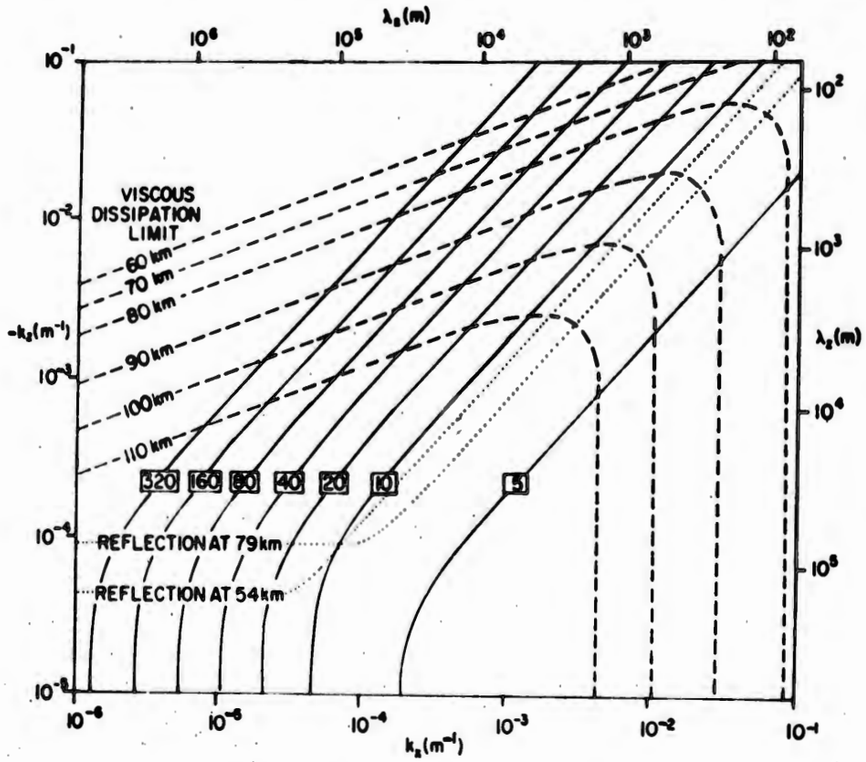

FIG. 11. Propagation modes at meteor heights, drawn for the same basic parameters as those of Fig. 9. The periods, measured in minutes, are shown in boxes on the corresponding constant-period contours (solid lines). The limits of the permitted spectrum, as determined by viscous damping, are shown for heights of 60, 70, 80, 90, 100, and 110 km (broken curves); modes lying above and to the right of these curves are excluded. The modes subject to reflection at heights of 54 and 79 km are also shown (dotted lines); modes lying below these curves cannot proceed from the lower atmosphere to the upper.

atmosphere (Minzner, Champion, and Pond 1959). They do not apply as well at intermediate levels where H is somewhat decreased, but the discrepancies are not serious.

The dissipative criterion (48) is illustrated by a family of curves drawn for heights of 60, 70, 80, 90, 100, and 110 km, or more precisely, for $\eta = 4.6\times10^{-2}$, 1.4×10^{-1}, 5.3×10^{-1}, 4.0, 36, and 3.0×10^{2} m²/s respectively (Minzner, Champion, and Pond 1959; the values up to 90 km are tabulated, the others derived from the governing formulae). Any one of these curves can be loosely interpreted as the boundary between the modes that can be propagated (which lie below on the diagram), and those that cannot (which lie above and to the right) at the height concerned. The progression of successive members of the family downward and to the left, as the height increases, illustrates the increasing importance of energy dissipation at greater heights. This results in part from the decrease of atmospheric density and in part, above 80 km, from the increase of temperature; both produce and increase in the kinematic viscosity, and lead to the extinction of an increasing number of propagation modes.

The parameters at a height of 225 km are sufficiently different from those at 100 km to warrant a separate presentation of the propagational characteristics; it is given in Fig. 12. The curves drawn there are based on an assumed $\gamma = 1.4$, and the values $g = 9.15$ m/s², $H = 51.3$ km (whence $C = 810$ m/s, $\tau_a = 13.2$ minutes, and $\tau_g = 14.7$ minutes), and $\eta = 2.4\times10^{5}$ m²/s, taken from Minzner, Champion, and Pond (1959). Figures 11 and 12 contain, in addition to the curves already discussed, others that result from considerations of reflection in the middle atmosphere. They will be interpreted shortly.

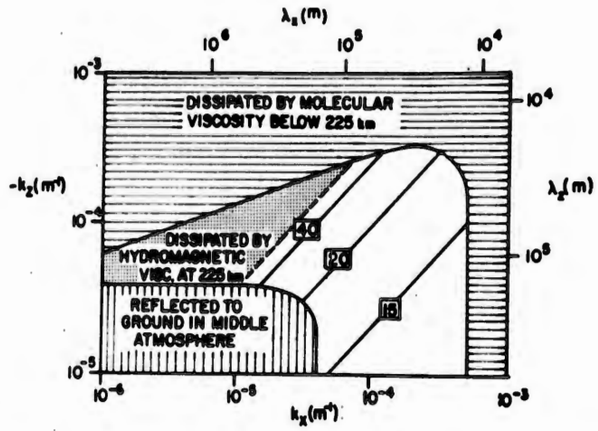

FIG. 12. Propagation modes at a height of 225 km ($\gamma = 1.40$, $g = 9.15$ m/s², $H = 51.3$ km; $\tau_a = 13.2$ minutes, $\tau_g = 14.7$ minutes; dH/dz assumed = 0). The periods, measured in minutes, are shown in boxes on the corresponding constant-period contours. Modes removed by reflection in the middle atmosphere and those removed by molecular viscosity in the upper atmosphere are represented by the hatched areas; those damped severely by hydromagnetic viscosity in the upper atmosphere (during daytime) are shown by the shaded area.

IV.5. Non-linear Effects

It may be recalled that a perturbation treatment was adopted for the manipulation of the basic equations. Such a treatment can be justified in application only by an a posteriori comparison of the relative magnitudes of the linear and non-linear terms, both being calculated by application of the relations derived in the linear treatment. The full equations appropriate to the general problem, and their perturbation-series expansions, are cumbersome and will not be dealt with in any detail here. An estimate of the importance of non-linear effects will be gained, instead, by a qualitative assessment.

It has already been noted that the atmospheric density undergoes fractional variations of 5–20% at some heights, and this indicates already that non-linearities cannot be ignored. The fractional changes of pressure are appreciably less—they are reduced by the ratio of the horizontal phase speed to the speed of sound, approximately, in the low-frequency asymptotic modes—and so add no non-linearities of any greater importance. The temperature and speed of sound will be subject to fractional variations comparable to those occurring in the density, but not appreciably greater.

The non-linearities of greatest potential importance appear to be those produced by the atmospheric motion. They arise in the mobile operator

$$D/Dt \equiv \partial/\partial t + \mathbf{U}.\,\mathrm{grad},$$

which should properly occur wherever $\partial/\partial t$ is to be found in equations (6)–(8). Their significance can be estimated from the ratio $|\mathbf{U}.\,\mathrm{grad}|/\omega$, which, for a single mode, becomes $|U_z k_z + U_x K_x|/\omega$. The largest individual terms between the numeric bars arise from the components $U_z k_z$ and $U_x k_x$, in the low-frequency asymptotic modes, but they cancel one another to a large extent because $\mathbf{U}$ is virtually perpendicular to the phase normals in those modes. The residual that does remain is quite significant nevertheless; it is given approximately by $(\gamma-1)^{-\frac{1}{2}} U_z/C$, and so exceeds the corresponding fractional variations of density by the factor $(\gamma-1)^{-1}$, or 2.5, to yield values as great as 50% self-modulation. The intermodulation between modes is even more important, for the cancellation just discussed no longer occurs; it results in relative non-linearities as great as $U_z/(\omega/k_z)$, and this ratio frequently is comparable to or even greater than unity.

From this brief assessment it will be evident that non-linear processes can be extremely important in the upper atmosphere, and indeed one of their possible consequences has already been given some attention in Section II.3. It seems likely that many of the relations employed in the analysis so far will be subject to quantitative error at the heights where these large non-linearities occur, but their qualitative validity is unlikely to be severely upset. In particular, the comparisons between theory and observation which were undertaken in Sections II and III are likely to remain significant despite the importance that is now being attributed to non-linear effects. This may be emphasized by noting that the calculated non-linearities apply only to a certain spectrum of modes, and only in a particular height range at that; those same modes at lower heights, and other modes, are much more amenable

to treatment by perturbation methods. (The intermodulation term, U. grad, cannot be wholly ignored in any mode at heights of 100–120 km, however, so long as some oscillatory motions comparable to ω/k_x are present.)

IV.6. Height Variations of Temperature

The analysis up to this point has been greatly simplified by the early assumption of an atmosphere which is uniform in both temperature and composition, or, for most practical purposes, uniform in the appropriate value of C or H. The more complicated case in which the speed of sound varies with height has been examined in some aspects by Martyn (1950), and it will evidently require further study in many more. The most important conclusions for present purposes can be derived by direct application of Martyn's work.

His equation (9′) can be rewritten as

$$\partial^2\phi/\partial z^2+H^{-1}(dH/dz)(\partial\phi/\partial z)+q^2\phi = 0 \tag{50}$$

where, in the present notation,

$$\phi = (\text{div } \mathbf{U}).\exp -\int(dz/2H) \tag{51}$$

$$q^2 = k_x^2(\omega_B^2/\omega^2-1)+(\omega^2-\omega_a^2)/C^2 \tag{52}$$

$$\omega_B^2 = (\gamma-1+\gamma dH/dz)g^2/C^2 \tag{53}$$

and z is the vertical Cartesian co-ordinate as before (and *not* the scaled co-ordinate employed by Martyn at this point); ω_B is the so-called Brunt frequency and $\tau_B \equiv 2\pi/\omega_B$ the Brunt period. A factor $\exp i(\omega t-k_x x)$ in div $\mathbf{U}$ is assumed but suppressed here, and the exponential growth of amplitude which has entered the earlier discussions is removed by the change of variable from div $\mathbf{U}$ to ϕ. Only terms of order dH/dz are added; those of order $(dH/dz)^2$ and d^2H/dz^2 are ignored. An approximate solution of the W.K.B. type can be found for (50), namely

$$\phi \propto |qH|^{-\frac{1}{2}}.\exp -i\int q\,dz. \tag{54}$$

In effect, then, (52) is the modified dispersion equation appropriate to height-varying conditions, with q^2 and ω_B^2 taking the parts previously played by k_z^2 and ω_a^2 respectively, when $q^2 > 0$. (It may be noted in passing that the derivation was in fact based on a varying C^2, not on a varying γ or g, so dH/dz should be interpreted as $dC^2/\gamma g dz$. Moreover, the result is applicable directly only to ϕ; U_x and U_z, for example, are individually subject to somewhat different dispersion relations.)

Martyn has interpreted (50) in a manner directly analogous to that adopted in the ray theory of radio-wave propagation through the ionosphere: wave energy progressing from a height where $q^2 > 0$, towards a height where $q^2 < 0$, will be reflected at the level where $q^2 = 0$. The application of this conclusion is of greatest immediate consequence at heights of 50–80 km, for there the value of q^2 decreases with increasing height in most modes of interest, and it can become negative for many of them. Modes in the latter category,

if generated low in the atmosphere, cannot penetrate to ionospheric heights (except with reduced energy flow, as deduced by a full-wave treatment; cf. the 12-hour solar tide).

The reflected modes are indicated in Fig. 11, by reflection curves drawn for heights of 54 km and 79 km. Each curve indicates the demarkation between reflected modes (below and to the right) and transmitted modes (above and to the left), and is determined by the condition $q^2 = 0$ at the height indicated. This condition has two asymptotic properties which lead to two distinct branches in the graphical representation: at low frequencies it approximates to

$$|\omega/k_x| \simeq C\omega_B/\omega_a, \tag{55}$$

which appears as a nearly horizontal line in Fig. 11, and at higher frequencies to

$$\omega \simeq \omega_B, \tag{56}$$

which appears as an inclined line. The heights of 54 and 79 km were chosen for representation because, on the model of the atmosphere adopted here (that of Minzner, Champion, and Pond 1959), ω_B has a minimum value ($\tau_g \simeq 8$ minutes) at 54 km and $C\omega_B/\omega_a$ has a minimum ($\simeq 170$ m/s) at 79 km. The illustrated curves then provide the extreme cases of reflection in the middle atmosphere, and the modes depicted by points lying above and to the left of both curves should escape serious reflection. The limits provided by the combination of these curves were employed earlier, in Fig. 6. The limit on ω/k_x provided by the 79-km height is the only one of interest at 225 km, since ω_g there is less than the value of ω_B at 54 km; it is depicted in Fig. 12 by a curve which is now horizontal only in part.

IV.7. Future Theoretical Analysis

Many courses of theoretical development are open to further exploration, and some have been indicated already. Four principal areas seem to stand out: (i) variations from the simple atmospheric model employed through the bulk of the foregoing discussions, (ii) the detailed effects of dissipative processes, (iii) the consequences of wave interaction, and (iv) the relation between the basic atmospheric motions and the consequent variations of electron density. These will be discussed briefly in turn.

(i) Some consequences of a change from an isothermal to a non-isothermal atmospheric model have already been pointed out. The reflection process to which this change gave rise has been discussed only briefly, and on the basis of what amounts to a ray treatment. The pertinent waves have, however, vertical wavelengths quite comparable to the scale of the temperature structure, and a full-wave treatment would therefore be more appropriate. This would lead to a picture in which partial reflection occurred for a broad range of modes, and total reflection in the middle atmosphere for none at all. The reflection curves of Figs. 11 and 12 would be broadened into a set of contours of varying transmission, and their form might be altered appreciably.

The possibility of ducting in a temperature-varying atmosphere should be explored along two further lines at least. Firstly, its significance to *F*-region travelling disturbances should be explored more fully, and some attention should be given to the corresponding energy density in the ducting region. The factor $q^{-\frac{1}{2}}$ in equation (54) indicates that a major ducted travelling disturbance in the *F* region might be accompanied by intense oscillations, and perhaps excessive turbulence, in the middle atmosphere where $q \simeq 0$. Secondly, the possibility of locally generated ducted modes should be explored, with particular reference to *E*-region heights. There the local τ_B ($\simeq 1$ minute) is appreciably smaller than in adjacent regions above and below, and oscillations at periods close to this value might very well be truly ducted. This possibility would be of importance in the normal ionosphere only if tidal generation were effective, but it could become of great significance during disturbed conditions when particle precipitation into the upper atmosphere is severe.

A further consequence of temperature variations can be of major importance, particularly when local generation is effective. It may be recalled that ω_a was necessarily greater than ω_g, and examination of the dispersion relation will show that this fact is of considerable importance in determining the pattern of τ-contours presented in Figs. 9–12. But, in a non-isothermal atmosphere, ω_B replaces ω_g in the dispersion equation and it is not necessarily less than ω_a. In fact, $\omega_B > \omega_a$ whenever $dH/dz > 0.064$ (if $\gamma = 1.40$), and this condition is satisfied both at heights of 25–45 km and again at all heights above 92 km (as determined on the basis of the model atmosphere of Minzner, Champion, and Pond 1959). The change is not appreciable for $\omega < 0.3\omega_a$, say, nor for $\omega > 3\omega_B$, but it becomes extreme when $\omega_a \lesssim \omega \lesssim \omega_B$. This is illustrated by Fig. 13, which is drawn for the conditions obtaining at a height of 110 km. Not only are the τ-contours radically altered, but the speeds of propagation are highly variable. The most drastic change of all, perhaps, is that which occurs in the direction and magnitude of energy or wave-packet propagation; this can reverse itself through infinite values along a single τ-contour, as is indicated. The wide variety of possibilities revealed by these considerations may provide some basis for an explanation of abnormal motions, particularly for those having extremely high apparent velocities.

Finally, it may be recalled that the possible variation of γ has been ignored. While this is appropriate up to *E*-region heights, its effect on *F*-region propagation should receive further examination.

(ii) Dissipative processes have received only casual treatment here, but their influence on the observable range of waves is certainly quite clear. They therefore warrant consideration in greater detail, with suitable revisions in the basic equations and with specific applications in mind. Each of the dissipation curves of Figs. 11 and 12 could be expanded into a set of contours indicating increasing rates of absorption or integrated absorption, and they might well be deformed in the process. In combination with contours of partial reflection, they should produce a family of propagational damping contours which might be fitted within the unshaded area of Fig. 6, for example. It is only in this way that selective propagational effects could be separated

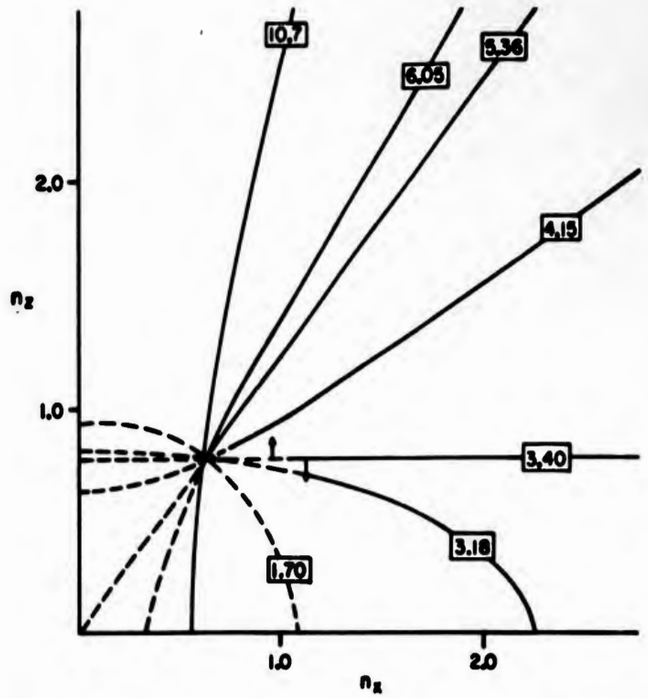

FIG. 13. Contours of constant period in the $n_x - n_z$ domain. The periods, measured in minutes, are shown in boxes on the corresponding contours. The basic parameters are $\gamma = 1.40$, $g = 9.5$ m/s^2, $H = 8.73$ km, and $dH/dz = 0.59$, whence $\tau_A = 5.36$ minutes and $\tau_B = 3.40$ minutes. These apply at a height of 110 km (Minzner, Champion, and Pond 1959). The dotted contours represent modes in which the vertical components of phase and energy progression lie in the same direction; the solid contours represent modes in which the vertical components of phase and energy progression lie in opposite directions. The transition from the dotted to the solid portion of a single contour takes place through infinite values of the energy velocity.

from selective generation, in an interpretation of the observed spectrum of wave modes. Turbulent and hydromagnetic viscosity must be explored, in addition to the molecular transport processes, in any exhaustive study.

The energy dissipated in the upper atmosphere by these waves merits study in relation to the general heat sources available there. Preliminary estimates indicate that the energy flux into the *F* region is normally an order of magnitude or more below that introduced by other processes (such as photoionization). The calculation involves much guessing as to the pertinent spectrum, however, and so cannot be considered as anything but a first attempt at an estimate. Moreover, it treats normal conditions only, and a significantly greater flux might well arise on individual occasions. The question of the wave contribution to the general heat balance must at the moment be considered an open one.

(iii) Wave interaction is another process that has received only cursory examination. If the earlier speculations concerning a cascade of energy are at all pertinent, then an understanding of the process is essential to a full interpretation of the spectral distribution of the observed modes. In any event, such an understanding is necessary if the energy of the dominant modes is to be traced through to its ultimate sink, for the non-linear terms

in the governing equations certainly become important for some modes before ordinary molecular dissipation becomes strong.

(iv) Most experimental studies of ionization irregularities in the E and F regions lack sufficient detail to warrant a thorough treatment of the perturbation motion of electrons, as distinct from that of the atmosphere proper. A notable exception to this statement has already been mentioned, however, concerning the formation of a sporadic-E layer by Dungey's process. A further exception has been implied: the interpretation of large-scale travelling disturbances will depend strongly on the interaction of the atmospheric motion and the geomagnetic field on the movement of ionization. As an extreme example, the effects of two atmospheric waves may be contrasted, the one having its perturbation velocity U lying along, and the second lying transverse to, geomagnetic field lines; the first can be expected to produce a far greater irregularity in the electron distribution than the second. Other extreme examples are provided by the resonance concepts previously discussed.

The actual motion of electrons will be affected by factors additional to those represented by these extreme cases. Their significance cannot yet be assessed, but it might very well lie, for example, in an explanation of seasonal variations in the observed directions of propagation (Munro 1958). Such variations could, of course, result from variations in the propagation directions of the atmospheric modes that happen to be generated. But equally, an observational selection could be operating through the enhancement or suppression of some of the atmospheric modes when these are transformed into irregularities of electron distribution, and the selection parameters may well vary with season. This suggestion is purely speculative, but it has the merit of bringing geomagnetic parameters into the selection process, and the preferred directions of travel do seem to be more closely related to geomagnetic than to geographic co-ordinates. In any event, an observational selection is certainly operative and its importance should be evaluated.

V. SUMMARY

Many properties of the irregular winds observed at meteor heights can be explained if it is assumed that the motion is produced by internal atmospheric gravity waves. These waves also provide an adequate explanation of 'drifts' in the distribution of ionization in the E and F regions, and of 'travelling disturbances' in the F region. All of these motions, and perhaps other manifestations of irregularities as well, can be fitted into a single comprehensive picture.

In one form, this picture envisages the generation of the waves in the energy-bearing regions of the lower atmosphere. A broad spectrum of waves may be anticipated, having independently varying directions of propagation, periods, and horizontal wavelengths. The vertical wavelength is determined by the period and horizontal wavelength, in a fashion that changes with height as the waves proceed through regions of varying temperature. Some of the upward-flowing energy is partially reflected in the middle atmosphere. The dominant modes that escape reflection attain maximum amplitude in the

E region, but there their energy is dissipated as they enter levels of ever-increasing kinematic viscosity. By the time the *F* region is reached, only a minor portion of the original spectrum remains. The amplitudes of oscillation are nevertheless large, for there is a continual increase of amplitude with height in all modes free from severe dissipation. Some of the energy that reaches the *F* region does so only after suffering a partial ducting in the middle atmosphere, and that ducting can carry the corresponding disturbances horizontally over very great distances with little loss of energy.

A more complete ducting could occur, and abnormal motions could arise, if the internal atmospheric gravity waves were generated locally in the *E* region. Non-linear interactions in the normal tidal oscillations might lead to this local generation, as could similar interactions in other waves propagating from the lower or middle atmosphere. Disturbances produced by, or in association with, the influx of particles at times of geomagnetic activity, provide an alternative local source of energy. Some evidence exists to suggest that this mechanism is indeed operative on occasion.

The further confirmation and elucidation of the wave theory appears to provide hope for an increased understanding of motions at ionospheric heights.

ACKNOWLEDGMENTS

The author is indebted to Dr. P. C. Clemmow and Dr. J. A. Fejer, who provided helpful suggestions, comments, and criticisms in the course of development of the foregoing theory. Discussions held with Dr. R. E. Barrington, Dr. J. S. Greenhow, Prof. K. M. Siegel, and many others, were also of great value. The work was undertaken under Project PCC D48-95-11-01 of the Defence Research Board of Canada; it was supported in part by the United States Air Force under Contract AF 19-(604)-5553, during the course of the author's consultancy at the Radiation Laboratory of the University of Michigan.

REFERENCES

AKASOFU, S.-I. 1956. Sci. Repts. Tôhoku Univ. Geophys. **8**, 24.
BIBL, K. and RAWER, K. 1959. J. Atmospheric and Terrest. Phys. **64**, 2232.*
BOOKER, H. G. 1955. J. Atmospheric and Terrest. Phys. **7**, 343.
——— 1956. J. Geophys. Research, **61**, 683.
——— 1959. J. Geophys. Research, **64**, 2164.*
BOWLES, K. L. 1959. J. Geophys. Research, **64**, 2071.*
BRIGGS, B. H. and SPENCER, M. 1954. Repts. Progr. in Phys. **17**, 245.
CHAPMAN, J. H. 1953. Can. J. Phys. **31**, 120.
DESSLER, A. J. 1959. J. Geophys. Research, **64**, 397.
DIEMINGER, W. 1955. The physics of the ionosphere (The Physical Society, London), pp. 53–57.
DUNGEY, J. W. 1955. The physics of the ionosphere (The Physical Society, London), pp. 229–236.
——— 1956. J. Atmospheric and Terrest. Phys. **8**, 39.
——— 1959. J. Geophys. Research, **64**, 2087, 2188.*
ECKART, C. H. 1960. Hydrodynamics of oceans and atmospheres (Pergamon, Oxford and New York).
FEJER, J. A. and VICE, R. W. 1959. J. Atmospheric and Terrest. Phys. **16**, 291.
GARDNER, F. F. and PAWSEY, J. L. 1953. J. Atmospheric and Terrest. Phys. **3**, 321.
GOSSARD, E. and MUNK, W. 1954. J. Meteorology, **11**, 259.
GREENHOW, J. S. 1959. J. Geophys. Research, **64**, 2085.*
GREENHOW, J. S. and NEUFELD, E. L. 1955. Phil. Mag. **46**, 549.
——— 1959. J. Geophys. Research, **64**, 2129.*

GREGORY, J. B. 1956. Australian J. Phys. **9**, 324.
HEISLER, L. H. 1958. Australian J. Phys. **11**, 79.
HINES, C. O. 1955. J. Atmospheric and Terrest. Phys. **7**, 14.
——— 1956. J. Atmospheric and Terrest. Phys. **9**, 56.
——— 1959*a*. J. Geophys. Research, **64**, 2210.*
——— 1959*b*. Proc. IRE, **47**, 176.
——— 1959*c*. J. Geophys. Research, **64**, 939.
JONES, I. L. 1958. J. Atmospheric and Terrest. Phys. **12**, 68.
LAMB, H. 1945. Hydrodynamics (Dover Publications, New York).
LILLER, W. and WHIPPLE, F. L. 1954. Special Supplement to J. Atmospheric and Terrest. Phys. **1**, 112.
MANNING, L. A. 1959. J. Geophys. Research, **64**, 1415.
MANNING, L. A., PETERSON, A. M., and VILLARD, O. G., JR. 1954. J. Geophys. Research, **59**, 47.
MARTYN, D. F. 1950. Proc. Roy. Soc. A, **201**, 216.
——— 1955. The physics of the ionosphere (The Physical Society, London), pp. 163–165.
——— 1959. J. Geophys. Research, **64**, 2043.*
MILLMAN, P. M. 1959. J. Geophys. Research, **64**, 2122.*
MINZNER, R. A., CHAMPION, K. S. W., and POND, H. L. 1959. The ARDC model atmosphere, 1959, Air Force Surveys in Geophysics No. 115, Air Force Cambridge Research Center, Bedford, Massachusetts.
MUNRO, G. H. 1950. Proc. Roy. Soc. A, **202**, 208.
——— 1953. Proc. Roy. Soc. A, **219**, 447.
——— 1958. Australian J. Phys. **11**, 91.
MUNRO, G. H. and HEISLER, L. H. 1956*a*. Australian J. Phys. **9**, 343.
——— 1956*b*. Australian J. Phys. **9**, 359.
PITTEWAY, M. L. V. 1958. Proc. Roy. Soc. A, **246**, 556.
RATCLIFFE, J. A. 1959*a*. J. Geophys. Research, **64**, 2102.
——— 1959*b*. J. Geophys. Research, **64**, 2057.*
STEWART, R. W. 1959. J. Geophys. Research, **64**, 2084.*

*These contributions are also published in the Proceedings of the International Symposium on Fluid Mechanics in the Ionosphere, Am. Geophys. Union, Washington, 1959.

UNCLASSIFIED